STUDIES OF PROTEIN-PROTEIN INTERACTIONS USING HUMAN RECOMBINANT NUCLEAR RECEPTOR CAR AND COACTIVATOR PGC1α

Maiju Piirainen

M.Sc. Thesis

Degree program in Biosciences, Major in Biochemistry

Faculty of Science and Forestry

University of Eastern Finland

October 2012

ABSTRACT

UNIVERSITY OF EASTERN FINLAND, Faculty of Science and Forestry

Degree program in Biosciences, Major in Biochemistry

Piirainen, Maiju Karoliina: Studies of protein-protein interactions using human recombinant nuclear receptor CAR and coactivator PGC1α

M.Sc. thesis 71 pages, 7 appendixes (78 pages)

Supervisors: PhD Ferdinand Molnár and PhD Sanna Ryhänen

October 2012

---

keywords: CAR, PGC1α, protein-protein interactions

Constitutive androstane receptor (CAR) is a nuclear receptor with classical function of mediation the metabolism of xenobiotics. However, a novel function has been discovered for CAR as a regulator of energy metabolism. It has been shown that CAR interacts with peroxisome proliferator-activated receptor γ coactivator 1 α (PGC1α), which is involved in various pathways connected to energy metabolism. PGC1α contains a conserved LXXLL motif and an arginine/serine-rich domain, and there is evidence that both regions interact with CAR.

The aim of this study was to study the interactions between these two proteins. Interactions were studied using full-length proteins in reporter gene assays and co-purification experiments were performed using the ligand binding domain of CAR and the LXXLL motif and RS domain of PGC1α. Reporter gene assay showed that PGC1α increases the activity of CAR by 5.7-fold. Co-purification showed that the RS domain of PGC1α seems to interact with CAR. However, the LXXLL motif, which is the conserved motif in coactivators important for the interaction with nuclear receptors, did not interact with CAR. However, this could be due to the low level of LXXLL expression in *Escherichia coli* (*E.coli*). The co-purification was also performed using synthetized LXXLL motif with optimized codons for the expression in bacterial cells, but the results were inconclusive. The co-purified protein should be confirmed with mass spectrometry. In conclusion, the activity of CAR is increased by PGC1α and there seems to be interaction between CAR and the RS domain of PGC1α. The LXXLL motif should be expressed in larger quantities in order to further investigate the interaction with CAR.

ACKNOWLEDGEMENTS

I would like to thank my supervisors Ferdinand Molnár and Sanna Ryhänen for their guidance and advice they have granted me. I would like to thank Professor Paavo Honkakoski for his support to my work in the research group. In addition, I would like to thank Viktória Prantner and Lea Pirskanen for their guidance in the laboratory. I am also grateful for the support of the whole group.

A huge thank you goes to my family and friends for the encouragements and support. Thank you for trying to understand what is it that I do.

Hämeenlinna 2.8.2012

Maiju Piirainen

ABBREVATIONS

aa	amino acid
AF	activation function
AhR	aryl hydrocarbon receptor
Akt/PKB	serine/threonine protein kinase/protein kinase B
AMPK	adenosine monophosphate-activated protein kinase
AR	androgen receptor
BAT	brown adipose tissue
bp	base pair
CAR	constitutive androstane receptor
cAMP	cyclic adenosine monophosphate
CITCO	6-(4-chlorophenyl)imidazo[2,1-b][1,3]thiazole-5-carbaldehyde O-(3,4-dichlorobenzyl)oxime
CREB	cAMP-responsive element binding protein
CYP	cytochrome P-450 monooxygenase
DBD	DNA-binding domain
E.coli	*Escherichia coli*
ER	estrogen receptor
ERR	estrogen-related receptor
FOXO1	forkhead transcription factor
FXR	farnesoid X receptor
GR	glucocorticoid receptor
Hat	histidine affinity tag
HAT	histone acetyltransferase
HNF4	hepatocyte nuclear factor 4
LB	Lennox L Broth Base
LBD	ligand-binding domain
LXXLL	leucine-X aa-X aa-leucine-leucine motif
LXXLLsyn	synthetic LXXLL motif
MCF-7	human caucasian breast adenocarsinoma
MR	mineralcorticoid receptor
NCoR	nuclear receptor corepressor

NR	nuclear receptor
OD_{600}	optical density at 600 nm
PB	phenobarbital
PBREM	Phenobarbital-responsive enhancer module
PCR	polymerase chain reaction
PCS	protein cleavage site
PEI	polyethyleneimine
PGC1	peroxisome proliferator-activated receptor gamma co-activator-1
PKA	cAMP-dependent protein kinase
PPAR	peroxisome proliferator-activated receptor
PR	progesterone receptor
PRC	PGC1 related coactivator
PXR	pregnane X receptor
RAR	retinoid acid receptor
RBS	ribosome binding site
RE	response element
RIP140	receptor interaction protein 140
RS	arginine/serine rich domain
RXR	retinoid X receptor
SB	super broth
SDS-PAGE	sodium dodecyl sulfate polyacrylamide gel electrophoresis
SHP	small heterodimeric partner
SIRT1	silent information regulator 2 homologue 1
SMILE	small heterodimer partner interacting leucine zipper protein
SMRT	silencing mediator of retinoid and thyroid hormone receptor
SRC-1	steroid receptor coactivator-1
T1DM	type 1 diabetes mellitus
T2D	type 2 diabetes
TCPOBOP	1,4-bis[2-(3,5-dichloropyridyloxy)]benzene
TH	thyroid hormone
TR	thyroid receptor
TRB-3	*tribbles* homologue 3
VDR	vitamin D receptor
WAT	white adipose tissue

TABLE OF CONTENTS

ABSTRACT 2
ACKNOWLEDGEMENTS 3
ABBREVATIONS 4
TABLE OF CONTENTS 6
1 INTRODUCTION 8
2 LITERATURE REVIEW 9
2.1 Transcriptional regulation 9
2.2 Nuclear receptors 10
2.2.1 Classification of NRs 10
2.2.2 Conserved modular structure of NRs 12
2.2.3 Binding of NRs to DNA 13
2.3 Xenobiotic NRs 14
2.3.1 CAR 14
2.3.2 CAR as a metabolic regulator 16
2.3.3 CAR and its connection to diabetes 18
2.4 Coregulators 19
2.4.1 Coactivators 20
2.4.2 Corepressors 22
2.5 PGC1 family of coactivators 23
2.5.1 Modular structure of PGC1α 23
2.5.2 Interaction of PGC1α with NRs 24
2.5.3 Energy homeostasis mediated by PGC1α 26
2.6 Interactions between CAR and PGC1α 29
2.7 Metabolic diseases 30
3 THE AIM OF THE WORK 32
4 MATERIALS AND METHODS 33
4.1 Reagents and media 33
4.2 Preparation of $RbCl_2$ competent cells 35
4.3 Comparison of human and mouse PGC1 homology using BLAST 35
4.4 Molecular cloning of the recombinant protein constructs 36

4.5 Analysis of the constructs 38
4.6 Mini scale protein expression tests 39
4.7 Western blot 41
4.8 Co-purification of hCAR and hPGC1α domains 42
4.9 Cell culturing and reporter gene assays 43
4.9.1 Optimization of reporter gene assay 43
5 RESULTS 45
5.1 Comparison of human and mouse PGC1α homology using BLAST 45
5.2 Analysis of the constructs 45
5.3 Mini scale protein expression tests 47
5.4 Western blot 49
5.5 Co-purification of hCAR and hPGC1α domains 51
5.6 Cell culturing and reporter gene assays 52
6 DISCUSSION 57
6.1 Comparison of human and mouse PGC1α homology using BLAST 57
6.2 Molecular cloning of the recombinant protein constructs 58
6.3 Analysis of the constructs 58
6.4 Mini scale protein expression tests 59
6.5 Western blot 60
6.6 Co-purification of hCAR and hPGC1α domains 61
6.7 Cell culturing and reporter gene assays 62
7 CONCLUSIONS AND FUTURE DIRECTIONS 63
8 REFERENCES 65
APPENDIX A 71
APPENDIX B 73

1 INTRODUCTION

Gene activation and repression are regulated with various factors, including nuclear receptors and coactivators. Active nuclear receptors bind to DNA and recruit other proteins of the transcriptional machinery for the gene activation. One nuclear receptor is constitutive andostane receptor (CAR), which classical function is to regulate the metabolism of xenobiotics, such as drugs and environmental contaminants (Honkakoski *et al.* 1998). Recent studies show that CAR is also connected to energy metabolism, for example the expression of CAR is induced by caloric restriction (Ding *et al*. 2006). In addition, CAR has been shown to down-regulate gluconeogenesis and lipogenesis by repressing factors inducing these processes (Gao *et al.* 2009, Dong *et al.* 2009). It has been shown that CAR is activated by peroxisome proliferator-activated receptor γ coactivator 1 α (PGC1α) (Shiraki *et al.* 2003). PGC1α is a protein regulating various energy pathways, such as induction of thermogenesis, mitochondrial biogenesis and gluconeogenesis. PGC1α is expressed abundantly in tissues with high energy consumption, such as brown adipose tissue, heart and skeletal muscle (Knutti *et al*. 2000).

The research on metabolic pathways is under great pressure due to the growing health problems of obesity and the onset of metabolic diseases, such as type 2 diabetes (T2D). This has generated the investigation of novel therapeutical targets. Both CAR and PGC1α have been connected to energy metabolism and could be possible targets. It is important to study the interactions of the proteins in the various pathways of energy metabolism for the design of new drugs.

There have been only few studies concerning the interactions of CAR and PGC1α using purified proteins. There is clear evidence that PGC1α coactivates CAR but the details of the interaction is still unclear. It has been published that the N-terminal LXXLL motif interacts with the ligand binding domain of CAR and also the C-terminal arginine/serine-rich (RS) domain is important for the activation (Shiraki *et al.* 2003). Further investigation is needed in order to specify the interaction regions in CAR and PGC1α complex.

2 LITERATURE REVIEW

2.1 Transcriptional regulation

Humans have approximately 20 000 – 25 000 genes and it is important that the regulation of these genes is under strict control (International Human Genome Sequencing Consortium 2004). The transcription is controlled by proteins including general transcription factors, co-factors, histones and chromatin remodeling proteins (Vaquerizas *et al.* 2009).

Efficient packaging of DNA is managed by wrapping 147 base pairs of DNA around eight histones forming a nucleosome and the DNA with proteins is referred to as chromatin. The tight association of DNA to histones represses the gene expression, because the binding sites for the coactivator complexes are inaccessible. Therefore the remodeling of the chromatin is an important step in gene activation. There are two classes of chromatin remodeling factors: ATP-dependent nucleosome remodeling complexes and factors containing histone acetyltransferase (HAT) activity (Glass & Rosenfeld 2000). ATP-dependent chromatin remodeling complexes, such as SNF2 protein family of ATPases, facilitate the exposure of DNA sequences from tight packaging to nucleosomes (Flaus & Owen-Hughes 2011). Histones are remodeled by post-translational modifications, such as acetylation, methylation, ubiquitination and phosphorylation. HATs acetylate the amino groups in the histone tails, which leads to opening of the tightly packed chromatin (Torok & Grant 2004).

The transcriptional machinery is a protein complex, which binds to gene regions and starts the transcription from DNA to RNA (Thomas & Chiang 2006). The preinitiation complex in eukaryotes includes RNA polymerase II (polII) and six additional transcription factors (TFIIA-H). TFIID binds to a TATA sequence, called TATA box, located in the promoter DNA and this event promotes the entry of other general transcription factors, such as TATA-binding protein-associated factors (TAFs) and the mediator protein complex. The TATA box is part of the core promoter region, which also includes the initiator and a downstream promoter element, that forms a platform for the transcriptional factors to bind.

The binding of the preinitiation complex is regulated by nuclear transcription factors including nuclear receptors (NRs) and other protein complexes regulating them. Upstream of the TATA box, there are proximal and distant enhancer regions, in which NRs and nuclear

factors can bind (Lemon & Tjian 2000). The factors binding to the enhancers can induce the gene transcription by recruiting coactivators with HAT activity and thus making chromatin accessible for preinitiaton complex, or by recruiting a kinase to phosphorylate the carboxy-terminal domain of polII and stimulate elongation (Farnham 2009).

2.2 Nuclear receptors

NRs are proteins, which function as transcriptional factors in the nucleus. NRs are involved in biological processes, such as development, cell differentiation and organ physiology, by modulating gene expression (Mangelsdorf *et al.* 1995). NRs function by interacting with different proteins forming complexes that repress or activate the gene expression. Many, but not all, NRs are activated by ligand binding, which results in recruitment of other proteins.

2.2.1 Classification of NRs

The human NR superfamily comprises from 48 receptors and they can be divided into six classes according to evolutionary analysis (Gronemeyer *et al.* 2004, Laudet 1997). The division of NRs to subfamilies is shown in Table 1. The first subfamily consists of receptors such as thyroid hormone receptor (TR), retinoic acid receptor (RAR), peroxisome proliferator-activated receptor (PPAR), vitamin D receptor (VDR), farnesoid X receptor (FXR) and constitutive androstane receptor (CAR). The second subfamily includes retinoic X receptor (RXR), hepatocyte nuclear factor 4 (HNF4) and some orphan receptors. The third subfamily consists of estrogen receptor (ER), estrogen-related receptor (ERR), glucocorticoid receptor (GR), androgen receptor (AR) and progesterone receptor (PR). The fourth subfamily comprises NGF-induced clone B, which is an orphan receptor and the fifth and sixth subfamilies include steroidogenic factor 1/Fushi Tarazu factor 1 and germ cell nuclear factor, respectively. There is also a subfamily 0 comprising small heterodimeric partner (SHP) and dosage-sensitive sex reversal (Laudet 1997, Aranda *et al.* 2001).

Table 1. Classification of NRs to subfamilies (Aranda & Pascual 2001, Gronemeyer *et al.* 2004, Cotnoir-White *et al.* 2010)

Subfamily	NRs	Name	Ligand
I	TRα, β	Thyroid hormone receptor	Thyroid hormone
	RARα, β, γ	Retinoic acid receptor	Retinoic acid
	PPARα, β, γ	Peroxisome proliferator -activated receptor	Fatty acids and their metabolites
	VDR	Vitamin D receptor	1,25-dihydroxy vitamin D_3, bile acids
	PXR	Pregnane X receptor	Xenobiotics, toxic endobiotics
	CAR	Constitutive androstane receptor	Xenobiotics, toxic endobiotics
	LXRα, β	Liver X receptor	Oxysterols
	FXR	Farnesoid X receptor	Bile acids
	RevErbα, β	Reverse ErbA	Heme
	RORα, β, γ	Retinoic acid-related orphan receptor	Cholesterol, cholesteryl sulphate, retinoic acid
II	RXRα, β, γ	Retinoid X receptor	Retinoic acid
	COUP-TFα, β	Chicken ovalbumin upstream prototer transcription factor	Fatty acids
	EAR-2	V-erbA-related protein 2	Orphan
	HNF4α, γ	Hepatocyte nuclear factor 4	Fatty acids
	TLX	Tailles-related receptor	Orphan
	PNR	Photoreceptor-specific nuclear receptor	Orphan
	TR2	Testis receptor 2	Orphan
	TR4	Testis receptor 4	Orphan
III	GR	Glucocorticoid receptor	Glucocorticoids
	AR	Androgen receptor	Androgens
	PR	Progesterone receptor	Progestins
	ERα, β	Estrogen receptor	Estradiol, tamoxifen
	ERRα, β, γ	Estrogen-related receptor	DES, 4-OH-tamoxifen
	MR	Mineralocorticoid receptor	Aldosterone, spirolactone
IV	NGFI-Bα, β, γ	NGF-induced clone B	Orphan
V	SF-1	Steroidogenic factor	Phospholipids
	LRH-1	Liver receptor homologue 1	Phospholipids
VI	GCNF	Germ cell nuclear factor	Orphan
0	SHP	Small heterodimeric partner	Orphan
	DAX-1	Dosage-sensitive sex reversal	Orphan

There are different ways to classify NRs to groups. Another way is to classify them according to their ligand properties and affinities into three groups: endocrine receptors, xenobiotic receptors and orphan receptors (Xie & Evans 2001, Sladek 2010, Li *et al.* 2003). Endocrine receptors can be referred as steroid hormone receptors or classical receptors and they bind ligands with high affinity (Gronemeyer *et al.* 2004). For example, endocrine receptors GR, ER, PR and VDR bind endogenous compounds glucocorticoids, estrogen, progesterone and vitamin D, respectively. Xenobiotic receptors were the first orphan receptors, but when

xenobiotic ligands were discovered, they were classified as environmental sensors (Sladek 2010). Thus, they are sometimes referred as adopted orphan receptors. Xenobiotic receptors bind ligands with lower affinity and broader specificity than endocrine receptors. CAR and PXR are classified into this group and they bind xenobiotics, such as drugs and environmental chemicals. Orphan receptors are true orphans for which no natural ligands have been identified and may be activated via alternative pathways (Xie & Evans 2001).

2.2.2 Conserved modular structure of NRs

The NRs have a similar overall structure composed of various functional domains (Aranda *et al.* 2001). The domains can be divided into five functional: A/B, C, D, E and F (Figure 1). The A/B region is in the N-terminus of the receptor and it contains the transcriptional activation function (AF-1) domain. The AF-1 domain mediates the ligand-independent activation of the receptor. The DNA-binding domain (DBD) is referred as C region and it was one of the first domain structures that had been discovered (Wrange & Gustafsson 1978). The DBD includes two regions called "zinc finger" consisting of four cysteines linked to one bivalent zinc ion. The D region acts as a hinge between the DBD and ligand-binding domain (LBD) and increases the flexibility of the structure. Some receptors have an additional F region at the C-terminus of the receptor but the precise function of this part is still unknown.

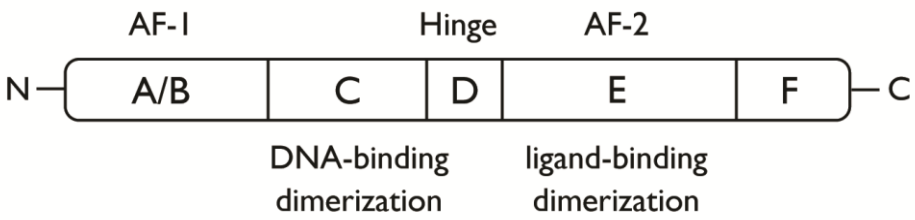


Figure 1. Schematic representation of the conserved NR structure. The A/B region contains the activation function domain 1 (AF-1). The E region contains the activation function domain 2 (AF-2), the ligand-binding pocket and the dimerization interface. The DNA-binding domain and the additional dimerization interface are located in the C region. The function of F region is unclear.

The LBD or E region positioned at the C-terminal part of the receptor has multiple functions in addition to the ligand-binding. It promotes dimerization of the receptors and the interactions with other protein partners. The NR LBD is comprised of 11-13 α helices and the three dimensional of the structural fold is well conserved for all NRs (Figure 2) (Wurtz *et al.*

1996). All the helices, except the very last one, form a ligand-binding pocket (domain) and the last helix (helix 12) that is part of the AF-2 acts as a lid for the LBD when the ligand is bound. The size of the binding pocket varies between NRs (Li *et al*. 2003). The helices 3, 4 and 12 have an important function for the interactions with cofactor proteins.

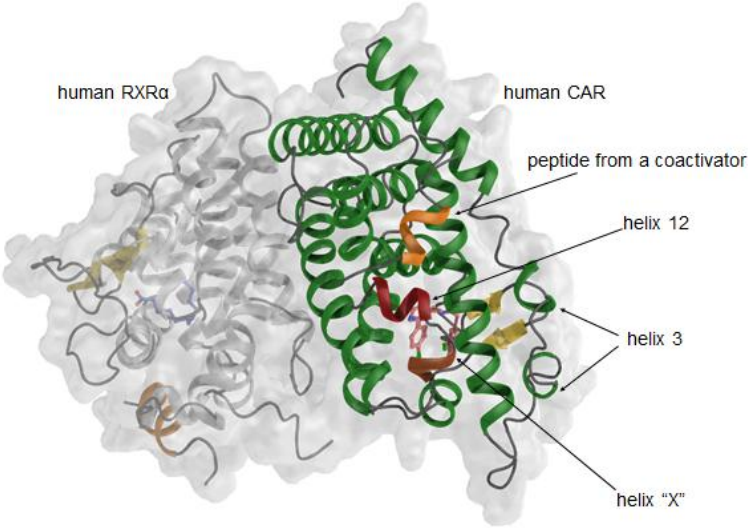


Figure 2. Crystal structure of the CAR LDB. There are twelve helices in the LBD and helices 3 and 12 are important for coactivator binding. Coactivator is shown in orange, helix 12 in red and the heterodimeric partner human RXRα in grey. Helix X is a short linker between helices 11 and 12.

2.2.3 Binding of NRs to DNA

NRs can function as monomers, homodimers or heterodimers when binding to the DNA (Gronemeyer *et al.* 2004, Aranda *et al.* 2001). Most the receptors in the subfamily 1 form heterodimers, however in the subfamily 2 and 3 monomers or homodimers are more common. RXR is a common heterodimerization partner for multiple receptors, including CAR, VDR and RAR (Lefebvre *et al.* 2010). The dimerization is crucial for high-affinity binding to the DNA. Both DBD and LBD structures have been shown to be involved in dimerization. The NRs bind to specific DNA sequences called response elements (RE), which have the consensus hexameric sequence RGKTCA (R = A or G, K = G or T) with three nucleotide spacing (Dunlop *et al*. 2004). REs are proximal or distal elements located around target genes and for dimeric NRs the two REs can be in different orientations called inverted repeats (IR), direct repeats (DR) or everted repeats (ER) and this characterizes the binding of NR dimers or monomers to DNA (Aranda & Pascual 2001, Frank *et al.* 2003, Helsen *et al.* 2012).

2.3 Xenobiotic NRs

Xenobiotics and drugs are metabolized in the liver to prevent the accumulation of potentially toxic substances in the body. There are three known receptors regulating the metabolism: aryl hydrocarbon receptor (AhR), constitutive androstane receptors (CAR) and pregnane X receptor (PXR) (Moreau *et al.* 2007). Elimination of the xenobiotics includes detection, metabolism, transport and excretion of the xenobiotic molecules. AhR, CAR and PXR act as xenosensors and regulate all the elimination processes mentioned above. AhR recognizes environmental substances including halogenated aromatic hydrocarbons, aromatic amines and polycyclic aromatic hydrocarbons, which are abundant in cigarette smoke and smog (Hankinson 1995). AHR belongs to a group of transcription factors, which contains conserved basic helix-loop-helix motif (Fukunaga *et al.* 1995).

Pregnane X receptor (PXR), referred as NR1I2 in the nomenclature system, was discovered in a screen of mouse liver library using expressed sequence tags (Kliewer *et al.* 1998). PXR is expressed abundantly in liver and small intestine and also weakly in kidney and stomach. PXR is activated by steroids, including pregnanes, such as pregnenolone-16α-carbonitrile (PCN), and it was alternatively referred as steroid and xenobiotic receptor, SXR, when discovered (Kliewer *et al.* 1998, Blumberg *et al.* 1998). Other ligands discovered to activate PXR are 5β-pregnane-3,20-dione, 17-OH-pregnenolone, RU486, cyproterone-accetate, clotrimazole, rifampicin and nifedipine (Quattrochi & Guzelian 2001).

2.3.1 CAR

Constitutive androstane receptor (CAR), also referred as NR1I3, is an adopted orphan NR, which is highly expressed in liver (Baes *et al.* 1994). CAR forms a heterodimer with RXR and was first discovered to bind to a retinoic acid response elements, which consists of hexameric DR sequences AGGTCA with five nucleotide spacing. CAR is constitutively active in the absence of a ligand. The activation occurs via helix 12 of the AF-2 motif, which mediates not only the ligand-dependent but also the ligand-independent transactivation of the receptor (Choi *et al.* 1997). The constitutive activity is based on the recruitment of transcriptional coactivators. The first ligands discovered for CAR were androstane metabolites, androstanol

and androstenol, but interestingly they act as inverse agonist and repress the activity of CAR (Forman *et al.* 1998).

CAR has a close relation to PXR since they both mediate drug metabolism by activating the gene expression of *CYP* genes (Honkakoski *et al.* 1998, Xie *et al.* 2004). CAR activates the expression of *CYP2B* genes in response to phenobarbital (PB) treatment. PB triggers the translocation of CAR from the cytoplasma to the nucleus, where CAR forms a heterodimer with RXR (Zelko *et al.* 2000). The CAR-RXR heterodimer binds to phenobarbital-responsive enhancer module (PBREM), which is located upstream of the *CYP2B* genes. Moreover, the CAR-RXR heterodimer can bind to XREM/ER-6 promoter sequences and enhance the activity of *CYP3A* genes.

In addition to PB, other ligands have been identified. Ligands may have different effects on CAR activity, for example agonists enhance and inverse agonists suppress the activity of CAR. The ligand recognition by CAR shows species-specific differences. The properties of some common ligands for human (hCAR) and mouse CAR (mCAR) are shown in Table 2. One of the ligands working as mCAR agonist is 1,4-bis[2-(3,5-dichloropyridyloxy)]benzene (TCPOBOP) (Tzameli *et al.* 2000). It selectively activates mCAR but has no effect on hCAR and has been shown to bind to hPXR (Moore *et al.* 2000). Another agonist, 6-(4-chlorophenyl)imidazo[2,1-b][1,3]thiazole-5-carbaldehyde O-(3,4-dichlorobenzyl)oxime (CITCO), is specific to hCAR (Maglich *et al.* 2003). In addition, ligands having an effect on CAR's activity and regulation of *CYP* genes, are chlorpromazine, 16-androstenes, progesterone and estrogens (Zelko *et al.* 2000). Clotrimazole binds to hCAR repressesing its activation (Moore *et al.* 2000).

Table 2. Common ligands binding to human and/or mouse CAR.

Ligand	human CAR	mouse CAR
Steroids:		
Androstenol	Strong inhibition	Weak inhibition
Estradiol	Inhibition	Activation
Xenobiotics:		
Phenobarbital	Activation	Activation
TCPOBOP	No effect	Strong activation
CITCO	Activation	No effect
Clotrimazole	Inhibition	No effect
Chlorpromazole	No effect	Activation

2.3.2 CAR as a metabolic regulator

In addition to drug and steroid metabolism, novel functions for CAR have been presented. It has been proposed that CAR plays a role in energy homeostasis and metabolism (Maglich *et al.* 2004). During caloric restriction CAR has been shown to decrease the thyroid hormone (TH), thyroxine and triiodothyronine, levels, which are important regulators of basal metabolic rate and are key modulators in weight loss and energy homeostasis on mammals (Kim 2008). There have been implications that CAR has various roles in induction, repression and blocking the expression of hepatic genes after PB treatment (Ueda *et al.* 2002). CAR represses the gene expression of enzymes important for energy metabolism including fatty acid oxidation and gluconeogenesis. In addition, CAR knockout mice showed induced gene expression of genes, such as *Cyp4a10* and *Cyp4a14*, which encode microsomal lipid peroxidases and thus stimulating the peroxidation. This indicates that blocking of lipid peroxisomes by CAR may protect from oxidative damage of lipids in liver.

The mechanisms of CAR-mediated response to caloric restriction crosstalks via multiple pathways (Ding *et al.* 2006). Fasting stimulates a pathway for the activation of CAR target genes, including *Cyp2b10*, *Sult2a1*, *Oatp2* and *Ugta1* (Figure 3). These genes are involved not only in drug metabolism but also in metabolism, transport and excretion of THs and other metabolites, such as steroids and bile acids. Fasting triggers a response in the body, where the levels of glucagon and adrenaline are increased. This activates the production of cyclic adenosine monophosphate (cAMP), which is an important secondary messenger in various biological processes. The increased intracellular level of cAMP activates PGC1α expression via cAMP-dependent protein kinase (PKA) and cAMP-response element binding protein (CREB). PGC1α coactivates HNF4α, which induces the expression of CAR. Finally, CAR and PGC1α complex induce the expression of CAR target genes by binding to the PBREM in the proximity of CAR target genes.

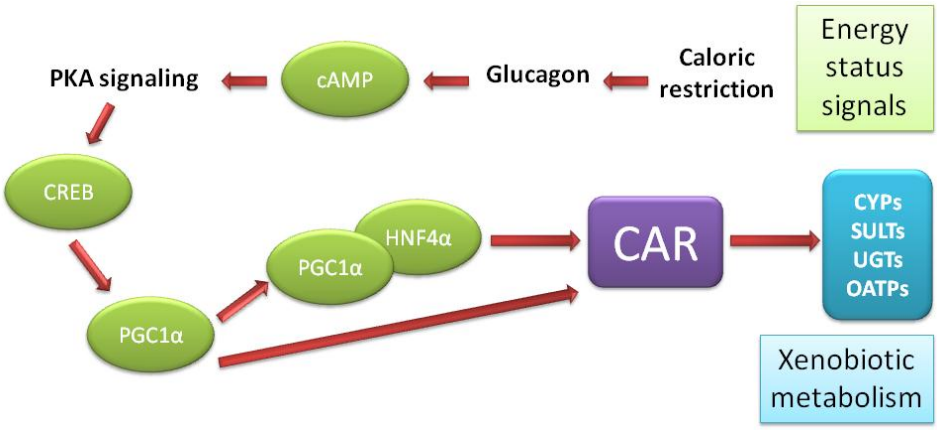


Figure 3. CAR mediated pathway in response to caloric restriction. Caloric restriction elevates the glucagon level, which increases cAMP. This results in activation of PKA signaling and CREB, which increases the expression of PGC1α. PGC1 α binds to HNF4 α activating the expression of CAR and also binds directly to CAR. This activation of CAR increases the expression of genes regulating the xenobiotic metabolism, which also regulates the metabolism of THs, other steroids and bile acids (modified from Ding *et al.* 2006).

It has been shown that CAR regulates the metabolism of serum triglycerides (Maglich *et al.* 2009). When CAR is induced with TCPOBOP treatment, the triglyceride levels increased during high-fat diet compared to control group. In CAR knockout mice, the serum triglyceride levels are at the same level as in wild-type mice without TCPOBOP treatment. The mechanism of CAR regulating the metabolism of triglycerides has been attributed to negative regulation of PPARγ activity. PPARγ regulates the gene expression of enzymes involving hepatic β-oxidation, which is an important pathway in fatty acid metabolism (Desvergne *et al.* 2006).

There have been contradictions in the effects of CAR with regard to weight loss. Some studies have shown that CAR increases the resistance to weight loss by decreasing the TH levels and therefore it has negative effects on obesity (Maglich *et al.* 2004, Ding *et al.* 2006). However, Gao and colleagues (Gao *et al.* 2009) showed that CAR is beneficial to obesity by inhibiting lipogenesis, secretion of very low density lipoprotein (VLDL), export of triglycerides and gluconeogenesis. In addition, CAR increases the energy expenditure of BAT and peripheral fat mobilization.

2.3.3 CAR and its connection to diabetes

Contrary to the effects on the resistance to weight loss, CAR has been shown to have positive effects on T2D symptoms (Dong *et al.* 2009). Activation of CAR decreases the serum glucose levels, represses gluconeogenesis and improves hepatic glucose uptake and insulin sensitivity. These effects are mediated by repression of gluconeogenetic genes phosphoenolpyruvate carboxykinase (PEPCK) and glucose-6-phosphatase (G6Pase) and by activation of glucose uptake genes for hexokinase and phophogluconate dehydrogenase (PGD). This pathway is mediated via CAR repressing the activity of forkhead transcription factor (FOXO1), which is an activator of PEPCK1 and G6Pase genes in the presence of TCPOBOP (Figure 4) (Kodama *et al.* 2004). The fact that CAR is activated by FOXO1 indicates that there is crosstalk between drug metabolism and energy metabolism. Insulin mediated signaling inactivates FOXO1 by exporting it from the nucleus to the cytoplasm and therefore downregulating CAR expression in the nucleus.

Repression of gluconeogenesis is not only mediated via FOXO1 pathway but there are also indications that CAR inhibits HNF4α (Figure 4) (Miao *et al.* 2006). CAR represses the activation of HNF4α in two mechanisms. Firstly, CAR-RXR competes with HNF4α for binding to the DR1 motif in the promoters of *CYP7A1*, *CYP8B1* and *PEPCK* genes thus inhibiting their gene expressions. Secondly, CAR competes with HNF4α by binding to common coactivators TIF-2 and PGC1α.

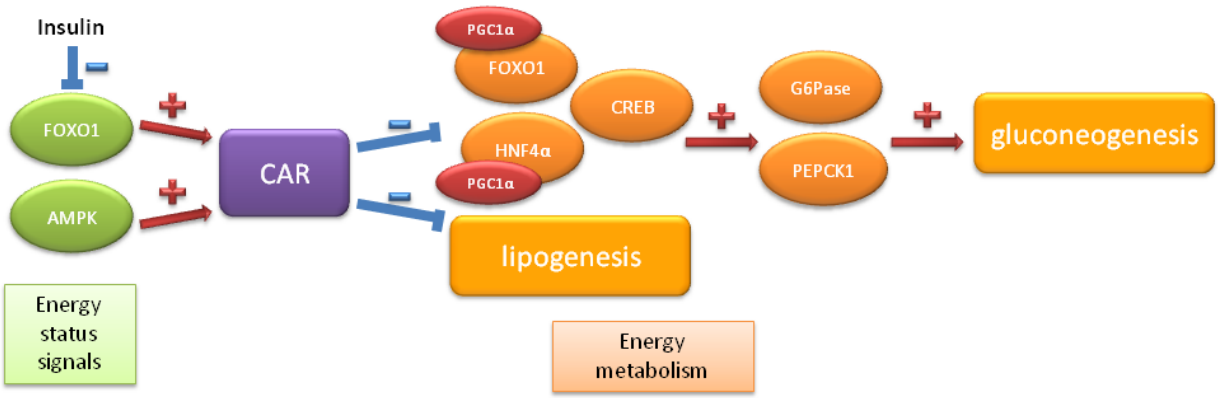


Figure 4. CAR mediated repression of factors responsible for energy metabolism. Energy status signals activate the expression of CAR, which represses lipogenesis and gluconeogenesis. CAR represses FOXO1, HNF4α and CREB, which increase gluconeogenesis by activating G6Pase and PEPCK1 (modified from Kodama *et al.* 2004, Moreau *et al.* 2007, Gao & Wie 2010).

CAR inhibits the expression of lipogenic genes in liver, WAT, BAT and skeletal muscle after treatment with TCPOBOP inhibited in high-fat diet mice (Gao *et al.* 2009). Due to the lack of CAR expression in WAT and BAT, it seems that these metabolic pathways are secondary to CAR activation in other tissues. It has been shown in the same study that PGC1α is suppressed by TCPOBOP treatment in this animal model, which is consistent with the metabolic benefit of CAR, contradicting the previous findings (Maglich *et al*. 2004).

CAR has been linked to the increased drug metabolism in type 1 diabetes mellitus (T1DM) (Dong *et al.* 2009). *Cyp2B10*, *Cyp3A11* and *Cyp2C29* gene expressions are elevated in T1DM mouse models and expressions of these genes is absent in CAR knockout mice. This indicates the importance of CAR in the elevated drug metabolism in untreated or poorly treated T1DM. The pathways for CAR's induction in T1DM were shown to involve elevated hepatic bile acids and adenosine monophosphate-activated protein kinase (AMPK) levels. Furthermore, the expression of PGC1α was increased indicating its contribution to the activation of CAR in T1DM.

2.4 Coregulators

The regulation of target genes by NRs requires the recruitment of other nuclear proteins (Rosenfeld & Glass 2001). One group of such proteins are coregulators that are divided to coactivators and corepressors, and are able to promote or repress the function of NRs, respectively. Binding of a ligand regulates the activation or repression of NRs and this promotes the recruitment of either coactivator or corepressor complexes, respectively (Rosenfeld *et al.* 2006).

Coregulators can be classified to five groups by their properties (Lemon & Tijan 2000). The first group consists of coactivators and corepressors, such as TAFs and TFIIA, which target the core transcriptional machinery and promoter recognition. Proteins in the second group regulate DNA-binding and interact with other coregulators and the transcriptional machinery. The third group consists of variable factors, which interact with RNA polymerase, other regulators and some of them have enzymatic activity. One example of a third class factor is the mediator, which is a multiprotein complex that interacts with RNA polymerase. The fourth group includes coregulators, which possess chromatin-modifying activities, such as acetyltransferases or deacetylases. The regulators can acetylate or deacetylate various targets,

for example histones, histone-related proteins, other activators and regulators and also the transcriptional machinery. The final group consists of ATP-dependent chromatin remodeling factors that utilize ATP in reactions, which modify the chromatin structure.

2.4.1 Coactivators

Coactivators can be divided into two groups, primary and secondary coactivators (Stallcup *et al.* 2003). Primary coactivators bind directly to NRs and recruit secondary coactivators and thus forming a transcriptional activation complex (Figure 5). The secondary coactivators act in ATP-dependent chromatin modeling complexes and have HAT activity. These coactivators can modify the chromatin structure by acetylation or methylation of histones or other proteins. The coactivation complex promotes the recruitment of RNA polymerase II complex by interacting with the components of the transcriptional machinery.

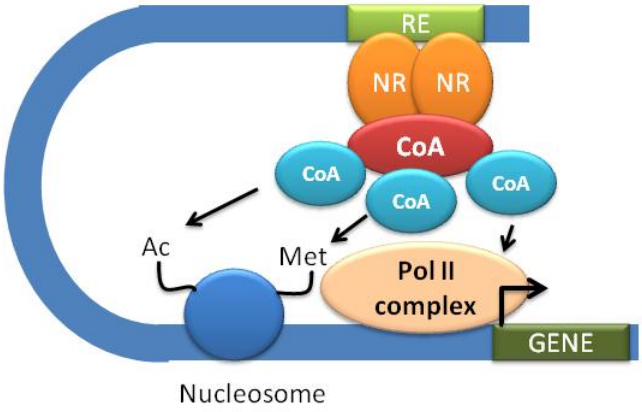


Figure 5. Primary and secondary coactivators. Primary coactivator shown in red, bind to NRs in orange and recruit secondary coactivators, shown in light blue. Secondary coactivators can acetylate or methylate histones in the nucleosomes thus activating transcription by helping the binding of the RNA polymerase II complex to DNA. Secondary coactivators can also directly interact with the transcriptional machinery (modified from Stallcup *et al.* 2003).

Coactivators have one or more special LXXLL (where L is leucine and X any amino acid) motif, which is also referred as NR box, and can interact with NRs (Heery *et al.* 1997). The amount of LXXLL motifs can vary and they can be located in any part of the coactivator protein. For instance, steroid receptor coactivator-1α (SRC-1α) has four LXXLL motifs, three located in the middle and one in the C-terminal part and both CREB-binding protein (CBP) and p300 have two in the N-terminal part. The LXXLL motif interacts with the helices 3, 4

and 12 all located in the LBD of NRs. The amino acid residues surrounding the core sequence of the LXXLL motifs are important for the binding affinity and specificity to NRs (MacInerney *et al.* 1998, Chang *et al.* 1999). The interaction between the LXXLL motif and NR LBD is mediated by hydrophobic interactions and a hydrogen bond between charged amino acid residues (Savkur & Burris 2004). A lysine residue in the helix 3 of NR LBD and a glutamic acid residue of helix 12 form the hydrogen bond with the LXXLL motif.

Coactivators have various roles in gene activation process. They can be involved in tissue promoter-specific activation, chromatin remodeling, platform proteins/HAT activation functions or mediator-like complex (Figure 6) (Rosenfeld & Glass 2001). One example of tissue promoter-spesific coactivator is PPARγ coactivator (PGC). Platform proteins function as interaction surface for other proteins to attach. They consist of the p160 family of NR coactivators, such as steroid receptor coactivator 1 (SRC-1)/nuclear receptor coactivator 1 (NCOA1), transcriptional intermediary factor 2 (TIF2)/GRIP1 and p300/CBP co-integrator associated protein (p/CIP)/amplified in breast cancer 1 (A1B1)/activator of the TR and RAR (ACTR)/receptor associated coactivator 3 (RAC3)/TR activator molecule 1 (TRAM-1) (Torchia *et al.* 1999). The p160 family interacts with the NR LBD and with secondary coactivators p300/CBP and methyltransferases, such as coactivator associated arginine methyltransferase 1 (CARM1) and protein arginine methyltransferase (PRMT1) via the AD2 domain (Chen *et al.* 1999, Koh *et al.* 2001). p300/CBP, CARM1, PRMT1 and p160 family have HAT activity and thus enhancing the transcription by modifying chromatin with acetylation of histones. SNF2 protein family of ATPases, including SWI/SNF and RSC, function as ATP-dependent chromatin remodeling regulators. ATPases hydrolyse ATP in reactions that are connected to modification of chromatin structure (Lemon & Tjian 2000).

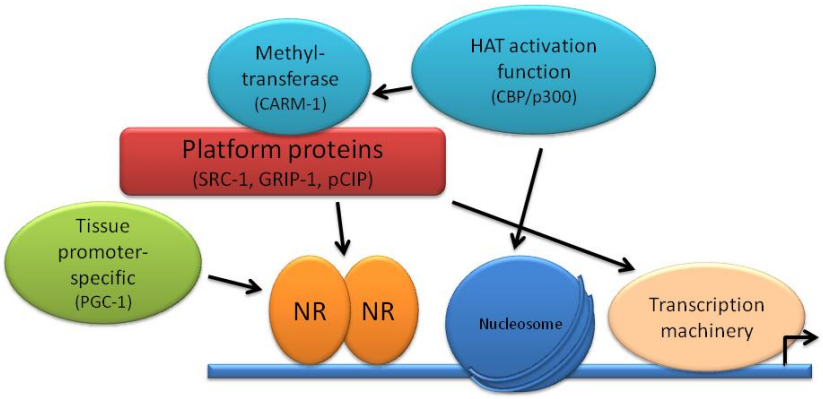


Figure 6. Schematic presentation of coactivator functions. Platform proteins bind other coactivators, such as CARM-1, but they also possess NR activation functions. Tissue promoter-specific proteins interact with NRs. Coactivators, which possess histone acetylationtransferase activation function modify the DNA packing to histone proteins. DNA and histone complex is referred as nucleosome. Acetylation of histones activates the binding of transcription machinery to gene initiation sequence. In addition, platform proteins can activate factors in the transcription machinery (modified from Rosenfeld &Glass 2001).

2.4.2 Corepressors

The other group of coregulators consists of corepressors, which are thought to repress gene transcription. The first corepressors discovered include nuclear receptor corepressor (NCoR) and silencing mediator of retinoid and thyroid hormone receptor (SMRT) (Xu *et al.* 1999). Other coregulators are receptor interacting protein 140 (RIP140), Hairless, Alien, ligand-dependent corepressor (LCoR) and small ubiquitous nuclear corepressor (SUN-CoR) (Lazar 2003, Aranda & Pascual 2001). RIP140 suppresses the expression of genes involved in metabolism (Christian et al 2006). It has four autonomous repression domains, but also contain has nine copies of the LXXLL motif, which is characteristic for coactivators and important for interaction with NRs (Christian *et al.* 2004).

The widely studied NCoR and SMRT have two important characteristics. First, they bind to NRs is ligand-independent and second, they contain autonomous, transferable repression domains (Lazar 2003). SMRT has two independent repressor domains (Nagy *et al.* 1997) and NCoR has three repressor domains (Heinzel *et al.* 1997). NCoR and SMRT contain similar sequences as LXXLL motif, also referred as the NR box, and copreressors interact with NRs by binding with these sequences, called corepressor/nuclear receptor (CoRNR) box (Hu & Lazar 1999). Both, NCoR and SMRT, function by forming a complex with mammalian Sin3 orthologue A (mSin3A) and histone deacetylase 1 (HDAC1) (Heinzel *et al.* 1997, Nagy *et al.*

1997). These complexes bind to NR in the absence of a ligand and HDAC1 forces the chromatin to inactive state and thus repressing the gene expression.

A novel corepressor, small heterodimer partner interacting leucine zipper protein (SMILE), has been identified as an inhibitor for ER, GR, CAR and HNF4α signaling (Xie *et al.* 2009). SMILE interacts with NR via the LBD and represses the activity of the NR by competing with coactivators, such as GRIP1 or PGC1α.

2.5 PGC1 family of coactivators

The PGC1 family of proteins consists of transcriptional coactivators, which recruits other transcriptional regulator proteins. The members of this have been shown to be involved in various metabolic pathways (Finck & Kelly 2006). The family has three subtypes: PGC1α, PGC1β and PGC1-related coactivator (PRC). Both PGC1α and PGC1β are found in tissues with high energy consumption and mitochondrial level, such as brown adipose tissue (BAT), skeletal muscle and heart. In addition, PGC1α is expressed in liver and kidney and lower levels in brain, lung and small intestine (Knutti *et al.* 2000). The biological function of PRC is less known. PGC1α was discovered in brown adipose tissue after cold-induction (Puigserver *et al.* 1998).

2.5.1 Modular structure of PGC1α

The N-terminus of PGC1α has a NR interaction motif, LKKLL (where L is leucine and K is lysine), at 144 to 148 amino acids (aa) (Knutti *et al.* 2000). There are also two other LXXLL (where X represents any amino acid) motifs in the N-terminus, LLAVL (88-92 aa) and LLKYL (210-214 aa) (Rha *et al.* 2009). The C-terminus of PGC1 contains two arginine-serine (RS)-rich domains (566-599 and 621-631 aa) and a RNA-binding domain (677-753 aa) (Knutti et al 2000). The schematic structure of PGC1α is shown in Figure 7.

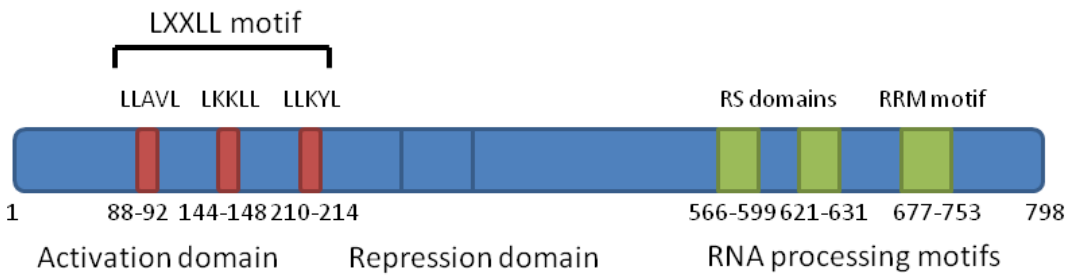


Figure 7 Schematic structure of PGC1α. PGC1α contains activation domain in the N-terminus and RNA processing motifs in the C-terminus. The LXXLL motifs are located in the activation domain and they are important for theinteraction with NRs. The RS domains and RNA recognition motif (RRM) are contributing to the RNA-processing (modified from Puigserver & Spiegelman 2003).

hPGC1α and hPGC1β have two conserved activation domains, AD1 and AD2, located in the N-terminal region (Sadana & Park 2007). There are five negatively charged aspartic and glutamic acid residues in these domains, and there is a glutamic acid residue at position 83, which is conserved between PGC1α and PGC1β. These amino acids were shown to be important for the transcriptional activation.

2.5.2 Interaction of PGC1α with NRs

PGC1α was originally discovered as a PPARγ binding protein (Puigserver *et al.* 1998). It was shown that the LXXLL motif in PGC1α and DNA binding domain in PPARγ are the regions that function as interaction surfaces. In addition, the deletion of the RS domain in PGC1α decreased the interaction and therefore it is involved in binding to PPARγ. More accurate information concerning the interaction of the LXXLL motif with PPARγ was discovered when it was shown that the LKKLL at residues 144-148 is sufficient for the interaction and activation by PGC1α (Li *et al.* 2008). The Puigserver group (Puigserver *et al.* 1998) also studied interactions with other NRs and TRβ, RARα and ERα bind to PGC1α. TRβ interacts with PGC1 in a ligand –independent manner, but ligand increases the binding by 2- to 3-fold. The binding of RARα and ERα to PGC1 seems to be ligand-dependent. More interactions between NRs and PGC1α were discovered when Knutti and coworkers (Knutti *et al.* 2000) investigated the activation of ER, AR, MR and GR by PGC1α in response to hormone addition. PGC1α enhanced the hormone-dependent transcription in ER, MR and GR and weakly in AR. Also ERRα was discovered to interact with PGC1α in heart tissue (Huss *et al.* 2002). The ERRα binding to PGC1α differs from other receptors. ERRα binds via LLKYL

motif at position 210-214 aa in PGC1α and other receptors bind to LKKLL motif at 144-148 aa.

Despite the primary interaction of one LXXLL motif, the other two motifs can have an intensifying effect to the activation and also the surrounding amino acids affect the binding properties (Rha *et al.* 2009). The LLAVL motif of PGC1α is surrounded by alanine and asparagine in the amino acids positions two (-2) and one (-1) prior to the first leucine, the LKKLL motif is surrounded by serine at position -2 and leucine at position -1 and the LLKYL is surrounded by serine at position -2 and glutamic acid at position -1 (Figure 8). Serine and asparagine have polar, uncharged side chains, alanine and leucine have hydrophobic side chains and glutamic acid has a negatively charged side chain. This could affect the binding properties of different LXXLL motifs to NRs.

```
                          -2-1 1 2 3 4 5+1+2
Position 79-99      DEENEANLLAVLTETLDSL
Position 137-155    EAEEPSLLKKLLLAPANTQ
Position 203-221    QRRPCSELLKYLTTNDDPP
```

Figure 8. Amino acids surrounding the LXXLL motifs in PGC1α. Three LXXLL motifs are indicated in red. Amino acid position from the first leucine in the LXXLL motif is numbered as 1, amino acids prior to the LXXLL motif are numbered as -1 and -2 and amino acids after the LXXLL motif are numbered as +1 and +2.

PGC1α as well as other proteins is part of the large signaling network regulating various pathways and processes in the cellular context. Some of the known and predicted interactions between PGC1α and other protein partners using the STRING 9.0 protein interaction database (http://string-db.org/) are shown in Figure 8. PGC1α is activated by SIRT1 and insulin, however there is evidence that insulin also inhibits PGC1α. Inhibition of PGC1α is mediated by FOXO1, which is inhibited by insulin. PGC1α activates EERα and UCP1, which is also activated by PPARγ. The evidence shows that PGC1α binds with FOXO1, ERRα, PPARγ, SIRT1 and CDK7.

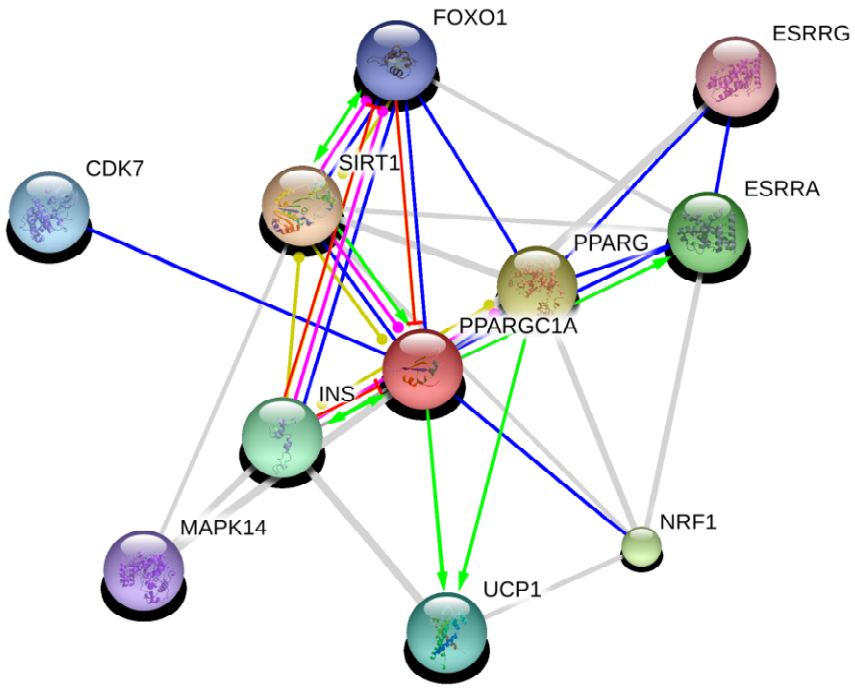


Figure 9. Interactions of PGC1α with NRs and other factors. Blue lines represent binding between the two proteins, green arrows represent activation, red lines represents inhibition, pink lines represent post-translational modifications, yellow lines represent expression and grey lines represent confidence of associations. SIRT = sirtuin 1, PPARG = peroxisome proliferators-activated receptor γ, FOXO1 = forkhead box O1, INS = insulin, ESRRA = estrogen-related receptor α, ESRRG = estrogen-related receptor γ, CDK7 = cyclin-dependent kinase 7, MAPK14 = mitogen-activated protein kinase 14, NRF1 = nuclear respiratory factor 1, UPC1 = uncoupling protein 1. Graph collected from STRING 9.0 online database for protein interactions (http://string-db.org/).

It has been shown that PGC1α regulates the metabolic pathways in the liver by binding to PXR and HNF4 (Bhalla *et al.* 2004, Jover *et al.* 2009). PXR suppresses HNF4α signaling of *cytochrome p450* (*CYP*) genes involving bile acid synthesis by competing for the binding of PGC1α.

2.5.3 Energy homeostasis mediated by PGC1α

When first discovered, PGC1α was shown to regulate the expression of uncoupling protein 1 (*UCP1*) (Puigserver *et al.* 1998). UCPs have an important role in thermogenesis and it is expressed in brown adipocytes and it functions in helping the oxidation processes to convert ADP to ATP in the inner mitochondria membrane (Flier & Lowell 1997). A recent study showed that PGC1α increases the expression of a membrane-bound protein, which is

modified into a hormone called irisin (Boström *et al.* 2012). Irisin stimulates the *UCP1* gene expression and brown-fat-like development. Irisin is increased with exercise and moreover, it increases the energy expenditure. Moreover, it has been shown that PGC1α activates mitochondrial biogenesis in muscle cells, BAT and heart (Wu *et al.* 1999, Kelly & Scarpulla 2004, Lehman *et al.* 2000). Muscle cells, adipocytes and heart tissue all have different levels of coupled and uncoupled mitochondrial respiration and the tissue-specificity is achieved via PGC1α. The pathway of biogenesis is mediated by transcriptional factors called nuclear respiratory factors (NRFs). Environmental stimuli, such as cold and high-calorie diet, release norepinephrine, activates β-adrenergic receptors in the cell surface. This results in an increase of cAMP in the cytoplasm and leads to the induction of *PGC1A* gene expression. PGC1α induces the expression of NRFs and functions as a coactivator of NRF-1 activating the expression of genes encoding subunits of mitochondrial respiratory chain and mitochondrial transcription factor A.

PGC1α is connected to mitochondrial respiratory processes by regulating reactive oxygen species (ROS), which are produced in the electron transport chain (St-Pierre *et al.* 2006). PGC1α is important to the oxidative stress defense system by inducing ROS-detoxifying enzymes, such as glutathione peroxidase, catalase and superoxide dismutase. This has notable results in the protection of neuronal cells, which are sensitive to the oxidative stress.

Moreover, both PGC1α and PGC1β are important for mitochondrial biogenesis in insulin resistant mouse heart (Mitra *et al.* 2012). Loss-of-function studies showed that mice lacking PGC1α have decreased respiratory function and altered mitochondrial gene expression, which are induced by high fat diet in wild-type mice. Therefore lipid metabolism and mitochondrial dysfunction have important roles in diabetic heart failure. However, it has been also shown that the effects of PGC1α for mitochondrial biogenesis are restricted in insulin resistance, glucose intolerance or early stages of diabetes and in the later stages of diabetes PGC1α expression is declined.

In addition to tissues with high-energy consumption, PGC1α is expressed in the liver and it is induced during fasting or caloric restriction (Yoon *et al.* 2001). cAMP and glucocorticoids levels are elevated in the liver by fasting and it has been showed that these molecules activate the gene expression of PGC1α. Moreover, PGC1α increases the synthesis of glucose by

stimulating the expression of gluconeogenic enzymes, such as PEPCK and G6pase. In addition, it has been shown that PGC1α furthers insulin resistance in the liver (Koo *et al.* 2004). PGC1 function with PPARα and activates mammalian *tribbles* homologue 3 (TRB-3), an inhibitor of the serine-threonine kinase/protein kinase B (Akt/PKB). Akt/PKB is part of insulin signaling pathway, which decreases the glucose output in the liver. TRB-3 is induced during fasting, in result of induction of PGC1 and PPARα. Tissue-specific functions of PGC1α are shown in Figure 10.

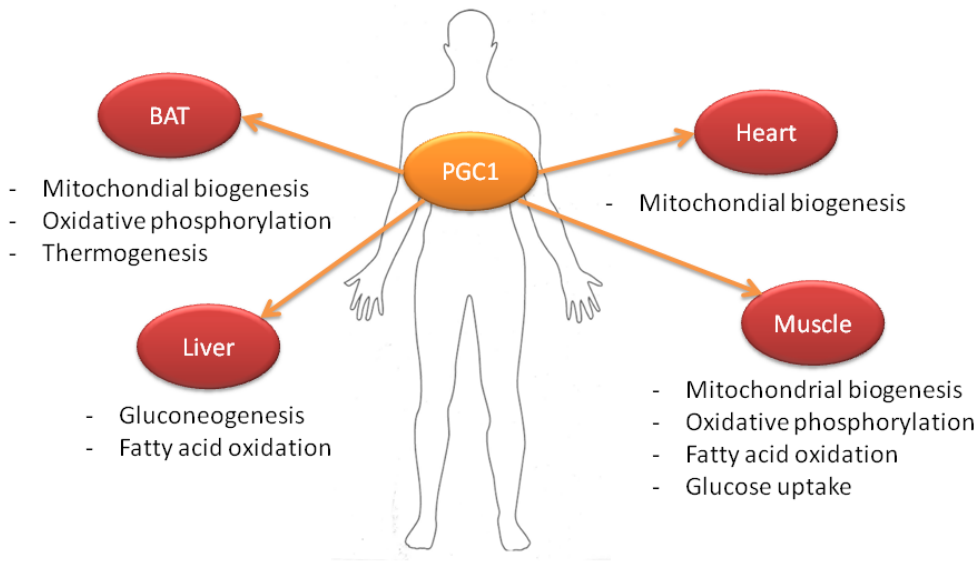


Figure 10. PGC1α and its functions in the human body. PGC1α has tissue-specific functions and it is expressed in brown adipose tissue (BAT), liver, heart tissue and skeletal muscle.

Bile acids are important regulators of cholesterol metabolism in the liver and their homeostasis is controlled by negative feedback through a NR, SHP (Eloranta & Kullak-Ublick 2004). Bile acids are also shown to act as ligands for various NRs. Bhalla and colleagues (Bhalla *et al.* 2004) showed that binding of PXR to PGC1α represses bile acid synthesis in the liver. However, it has been discovered that bile acids inhibit PGC1α expression via SHP (Yamagata *et al.* 2007). Bile acids bind to FXR and this promotes the expression of SHP, which then interacts with FOXO1 and the complex binds into the promoter of *PGC1α* gene repressing the expression.

PGC1α regulates triglyceride metabolism by activating FXR (Zhang *et al.* 2004). The mRNA transcription of *PGC1α* is increased during fasting and it increases the *FXR* transcription by coactivating PPARγ and HNF4α. PGC1α-PPARγ and PGC1α-HNF4α complexes bind to

DR-1 element located in the FXR promoter region. Moreover, PGC1α binds directly to FXR to activate the FXR target genes that up-regulates the triglyceride clearance and down-regulates triglyceride synthesis and leading to the activation of the fatty acid β-oxidation.

Due to the positive effects of PGC1α, there have been attempts to increase the levels of PGC1α. Exercise has been shown to elevate PGC1α levels (Franks & Loos 2006). In addition, food substances can induce PGC1α expression. Resveratrol, a polyphenolic compound found in the skin of the grapes and thus in red wine, was discovered to increase the activity of PGC1α (Lagouge *et al.* 2006). Resveratrol can induce the gene expression of PGC1α and moreover, regulate the post-translational modifications, for example acetylation. In this case, resveratrol indirectly activates silent information regulator 2 homologue 1 (also called sirtuin 1, SIRT1), which then deacetylates and hence activates PGC1α. Sirtuins are proteins, which possess NAD^+-dependent deacetylase and/or ADP-ribosyltransferase activities, and they have been connected to aging and energy metabolism (Li & Kazgan 2011). SIRT1 functions with PGC1α in liver and in adipose tissue.

2.6 Interactions between CAR and PGC1α

There have been only a few studies done concerning the interactions between CAR and PCG-1α. These interactions have been studied using reporter gene analysis. It has been showed that CAR competes with HNF4α for binding to PGC1α, which results in repression of the HNF4α target genes (Miao *et al.* 2006). Interactions between CAR and PGC1α were shown in a study, in which corepressor SMILE competes with PGC1α and this decreases the activation of CAR (Xie *et al.* 2009). Moreover, SMILE declines the expression of CAR target gene *CYP2B6.*

The details of the interaction between CAR and PGC1α have been described by Shiraki and co-workers (Shiraki *et al.* 2003). They have published that CAR is activated by PCG-1α in a ligand-independed manner and have showed that the RS domain of PGC1α targets CAR to nuclear speckles but both regions, LXXLL and RS, are interacting with CAR. Deletions in PGC1α has showed that the removal of the LXXLL motif decreases CAR's activity but does not abolish it and the removal of both RS and RRM domains also decreases its activity. Thus both regions, LXXLL and RS, are important for the activation of CAR. A schematic

representation of the target gene activation by interaction of PGC1α and CAR is shown in Figure 11.

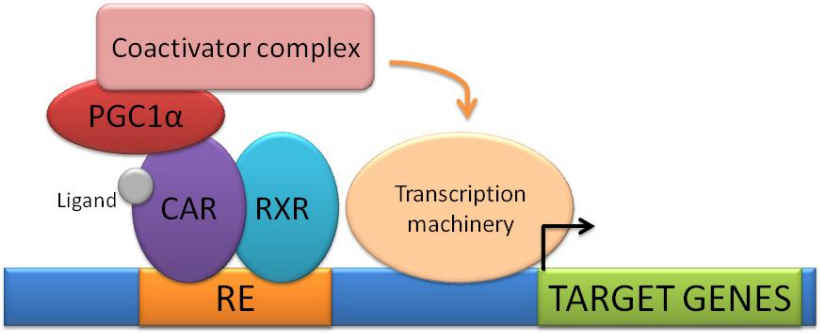


Figure 11. Gene activation by PGC1α and CAR interaction. PGC1α interacts with CAR and the CAR-RXR heterodimer complex binds to the response element (RE) in the promoter DNA of the target gene. PGC1α recruits other coactivators, which activate the transcriptional machinery to initiate the expression of the target gene (modified from Finck & Kelly 2006).

2.7 Metabolic diseases

Metabolic diseases including obesity and cardiovascular disease related factors, such as high blood pressure, high levels of tricylglycerols and inflammation are increasing in the western population (Guarente 2006). These physiological symptoms are referred as metabolic syndrome. Metabolic syndrome is also connected to dysregulation of glucose homeostasis, which can lead to T2D.

Obesity is a global health risk factor and in the year 2005 there were approximately 940 million overweight and 400 million obese adults in the world (Kelly *et al.* 2008). Overweight and obesity are defined as body mass index 25-29.9 kg/m^2 and ≥ 31 kg/m^2, respectively. It has been estimated that in the year 2030, there will be 1.35 billion overweight and 573 million obese adults. This fact has outstanding effect on the health of the world's population and has economic aspects due to other health issues it precedes, such as cardiovascular disease, cancer, diabetes and other associated complications.

There are two types of diabetes mellitus. T1DM is known as insulin-dependent diabetes mellitus and it is an autoimmune disease and usually onsets at an early age (Fowler 2010). The second type is T2D, which is non-insulin-dependent and is connected to metabolic

dysfunction. T2D is caused by combination of insulin resistance, β-cell dysfunction and increased hepatic glucose production. Insulin resistance is associated with obesity and may be one of the reasons why T2D is common in obese individuals. However, not all obese individuals develop T2D. The crucial factor is β-cell dysfunction, which leads to decreased insulin release (Kahn *et al.* 2006). This affects the glucose regulation by increasing hepatic glucose production and reducing glucose uptake in insulin-sensitive tissues. Despite the decrease of insulin release in dysfunctional β-cells, the insulin resistance is compensated with enhanced β-cell mass, which leads to hyperinsulinemia (Polonsky 2000).

Due to the extent public health related and economic effects, there is continuous research in finding therapeutical targets for obesity, metabolic syndrome and T2D. One investigated approach is targeting the transcriptional regulation complexes involved in the gene expression of the mediators in the metabolic pathways.

3 THE AIM OF THE WORK

The aim of this practical part was to investigate the interactions of different regions of hPGC1α and hCAR-LBD. The work also includes the preparation of the tools necessary for the experiments. The aims can be summarized into the following points:

1. Preparation of the co-expression vectors containing the recombinant hCAR-LBD and hPGC1α-LXXLL, hPGC1α-RS and synthetic hPGC1α-LXXLL domains.
2. Full optimization of the constructs using various *E.coli* bacterial host strains.
3. Characterize the interaction between hCAR-LBD and hPGC1α domains using co-purification.
4. Study the activation of full-length CAR by full-length hPGC1α using reporter gene assay in mammalian cellular models.

4 MATERIALS AND METHODS

4.1 Reagents and media

Table 3. List of chemicals.

Name	Company
Carbenicillin	Carbenicillin DIRECT, UK
Chloramphenicol	Sigma-Aldrich Co. LLC., St. Louis, USA
Isopropyl-β-D-1-thiogalactopyranoside (IPTG)	Fermentas, Thermo Fisher Scientific Inc., Waltham, USA
L-rhamnose monohydrate	Carbosynth Limited, Berkshire, UK

Table 4. List of bacterial growth media.

Name	Content	Company
Lennox L Broth Base (LB)	1 % peptone 0.5 % yeast extract 0.5 % NaCl	BD, Franklin Lakes, USA BD, Franklin Lakes, USA Sigma-Aldrich Co. LLC., St. Louis, USA
Super Broth (SB)	3.2 % tryptone 2 % yeast extract 0.5 % NaCl	BD, Franklin Lakes, USA BD, Franklin Lakes, USA Sigma-Aldrich Co. LLC., St. Louis, USA
N-Z-amine	1 % N-Z-amine A® from bovine milk 0.5 % yeast extract	Sigma-Aldrich Co. LLC., St. Louis, USA BD, Franklin Lakes, USA

Table 5. List of *E. coli* strains used for cloning and expression.

Strain	Usage	Company
XL-10	DNA amplification	Stratagene, Agilent Technologies, Inc., Santa Clara, USA
BL21(DE3)	Protein expression	Novagen EMD Millipore, Billerica, USA
BL21-Gold (DE3)	DNA amplification and protein expression	Stratagene, Agilent Technologies, Inc., Santa Clara, USA
BL21(DE3) pRARE	Protein production	Novagen EMD Millipore, Billerica, USA
Origami B (DE3)	Protein production	Novagen EMD Millipore, Billerica, USA

Table 6. List of mammalian cell lines.

Name	Abbreviation	Source	Recommended medium
Human hepatocellular carcinoma	C3A	ATCC, LGC Standards AB, Boras, Sweden	ATCC-formulated Eagle's minimum essential medium, 10 % fetal bovine serum
Human caucasian breast adenocarcinoma	MCF-7	ATCC, LGC Standards AB, Boras, Sweden	ATCC-formulated Eagle's minimum essential medium, 0.01 mg/ml bovine insulin, 10 % fetal bovine serum

Table 7. List of media for mammalian cells

Name	Content	Company
Sub-culturing medium	Dulbecco's Modified Eagle Medium with phenol red, 10 % fetal bovine serum, 1% L-glutamine, 1 % penicillin-streptomycin	Gibco®, Life Technologies Corporation, Carlsbad, USA
Sub-culturing medium before transfection	Phenol red-free Dulbecco's Modified Eagle Medium, 10 % fetal bovine serum, 1% L-glutamine, 1 % penicillin-streptomycin	Gibco®, Life Technologies Corporation, Carlsbad, USA
Transfection medium	Phenol red-free Dulbecco's Modified Eagle Medium, 1% L-glutamine, 1 % penicillin-streptomycin	Gibco®, Life Technologies Corporation, Carlsbad, USA
Ligand medium	Phenol red-free Dulbecco's Modified Eagle Medium, 5 % lipid-free serum, 1% L-glutamine, 1 % penicillin-streptomycin	Gibco®, Life Technologies Corporation, Carlsbad, USA

4.2 Preparation of $RbCl_2$ competent cells

Competent *E. coli* BL21 cells were prepared using Rubidium method. Cells were inoculated from a glycerol stock and grown in 5 ml LB medium overnight at 37 °C with 200 rpm shaking. Pre-culture was transferred to 250 ml LB and grown at 37 °C with 200 rpm shaking for 1 h 50 min, when optical density at 600 nm (OD_{600}) was measured with spectrophotometer (Varian Cary 50 UV/Vis spectrophotometer, Agilent, Santa Clara, USA). The cultivation was continued until OD_{600} reached 0.4. Then the culture was collected using 50 ml falcons and incubated for 15 min on ice. Cells were harvested by centrifugation at 2500 rpm for 15 min at 4 °C. Pellets were resuspended in 10 ml of transformation buffer 1 (TFB1: 10 mM MES, 100 mM RbCl, 10 mM $CaCl_2 \bullet 2\ H_2O$, 50 mM $MnCl_2 \bullet 2\ H_2O$, pH 5.8), then split into two 50 ml falcons and the volume was increased to 40 ml with TFB1 in both falcons. Resuspended mixtures were incubated on ice for 15 min and centrifuged at 2500 rpm for 15 min at 4 °C. Supernatants were discarded, the pellets were resuspended in transformation buffer 2 (TFB2: 10 mM MOPS, 10 mM RbCl, 75 mM $CaCl_2 \bullet 2\ H_2O$, 15 % glycerol, pH 6.5) and incubated on ice for 15 min. The suspension of cells was aliquoted in microcentrifuge tubes, flash frozen in liquid nitrogen and stored at – 85 °C until further use.

4.3 Comparison of human and mouse PGC1α homology using BLAST

Comparisons of nucleotide and amino acid sequences were performed in order to clarify the similarity between human and mouse PGC1α. Nucleotide sequences were acquired from National Center for Biotechnology Information (NCBI) Nucleotide database (http://www.ncbi.nlm.nih.gov/nuccore) and nucleotide sequences were translated to amino acid sequences using a translate tool at Bioinformatics Resource Portal called ExPASy, at Swiss Institute of Bioinformatics (SIB) web site (http://web.expasy.org/translate/). The comparison between human and mouse PGC1α was performed using on-line Basic local alignment search tool (BLAST) at NCBI web site. Alignment was performed for both, nucleotide and amino acid, sequences using nucleotide blast and protein blast, respectively.

4.4 Molecular cloning of the recombinant protein constructs

hCAR-LBD and hPGC1α-motifs/domains were cloned into a co-expression vector , pCoAT7H3F, containing a histidine affinity tag HAT (Takara Bio Europe/Clontech, Saint-Germain-en-Laye, France) for the first protein coding gene and three copies of a synthetic FLAG protein tag (Sigma-Aldrich Co. LLC., St. Louis, USA) for the second gene. The vector contains ampicillin resistance, T7 promoter and terminator, *lac* operator (*lac*O), TEV (first gene) and 3C protein (second gene) cleavage sites (PCS), ribosome binding sites (RBS) and gene for *lac* repressor protein (*lac*I). T7 promoter and *lac*O are located in the 5' promoter of both cloned protein genes. PCS are located after the affinity tags in order to facilitate removal of the tags if necessary. Table 8 summarizes the classical cloning of the inserts into co-expression vector. In addition, a synthetized LXXLL domain (LXXLLsyn) with bacterial codon optimization was cloned into the pCoAT7H3F. LXXLLsyn domain was created for better expression of the recombinant protein in *E. coli* due to the bacterial codon optimization but its size is smaller compared to the LXXLL motif created from the wild-type full-length PGC1α. The maps of the constructed plasmids are shown in the appendix A.

Table 8. Constructs prepared for recombinant protein expression and co-expression.

Construct #	Insert				Vector
	Hat-hCAR		3xFLAG-hPGC1α		
	Domain	REs	Domain	REs	
1	LBD	*Bam*HI/*Nde*I			pCoAT7H3F
2	LBD	*Bam*HI/*Nde*I	FL	*Bam*HI/*Nco*I & *Nco*I	pCoAT7H3F
3	LBD	*Bam*HI/*Nde*I	LXXLL	*Nco*I	pCoAT7H3F
4	LBD	*Bam*HI/*Nde*I	RS	*Bam*HI/*Nco*I	pCoAT7H3F
5	LBD	*Bam*HI/*Nde*I	synthetic LXXLL	*Bam*HI/*Nco*I	pCoAT7H3F
6	-	-	FL	*Bam*HI/*Nco*I & *Nco*I	pCoAT7H3F
7	-	-	LXXLL	*Nco*I	pCoAT7H3F
8	-	-	RS	*Bam*HI/*Nco*I	pCoAT7H3F
9	-	-	synthetic LXXLL	*Bam*HI/*Nco*I	pCoAT7H3F

The full-length hPGC1α was amplified with polymerase chain reaction (PCR) with Phusion polymerase (Finnzymes, Thermo Fisher Scientific Inc., Waltham, USA) from cDNA template. PCR conditions are shown in Table 9 and the primers used are shown in Table 10. The PCR product was purified using PCR purification kit (Qiagen, Hilden, Germany). The purified PCR product was digested with *Nco*I alone or *Nco*I/*Bam*HI in order to create different hPGC1α domain fragments such as hPGC1α-LXXLL was created with *Nco*I and hPGC1α-RS with *Nco*I/*Bam*HI. The inserts were ligated into pCoAT7H3F after the 3xFLAG site. The ligation of the *NcoI*-cut LXXLL insert into pCoAT7H3F-hPGC1α-RS construct created the full-length hPGCa. hCAR-LBD was digested from pET15b-hCAR$_{103-348}$ construct with *Nde*I/*Bam*HI and ligated into the empty pCoAT7H3F after the HAT site and also into the previously created pCoAT7H3F-hPGC1α constructs.

Table 9. PCR mixture for the amplification of the full-length hPGC1α.

Reagent	Amount	Cycle/Temperature		Time
DNA (50 ng/µl)	2 µl	98 °C		30 s
Buffer HF (5X)	40 µl	98 °C		7 s
dNTP (10 mM)	4 µl	55 °C	40 x	20 s
Primers (10 µM)	10 µl	72 °C		30 s
Phusion polymerase (2U/µl)	2 µl	72 °C		2 min
mQH_2O	142 µl	72 °C		∞
Total volume	200 µl			

Table 10. Primers used for PCR amplification of the full-length hPGC1α.

Primers for hPGC1α-FL amplification (5'-3'):	
hPGC1a_FL_BamHI_m2hF	ATTGGATCCATGGCGTGGGACATGGCA
hPGC1a_FL_CPEC_m2hR	CCCCAAGGGGTTATGGTAGGATCCTTACCTGCGCAAGCTTCT

Construct containing the N-terminal 3xFLAG-fusion of the synthetic hPGC1α-LXXLL domain was prepared by subcloning from pUC57-hPGC1α-LXXLLsyn plasmid obtained from BioBasic (BioBasic Inc. Unionville, Canada) into pCoAT7H3F vector after the 3xFLAG site using *Nco*I/*Bam*HI restriction enzymes. hCAR-LBD was inserted into the pCoAT7H3F-hPGC1α-LXXLLsyn plasmid as previously described for the other constructs.

Inserts were ligated into pCoAT7H3F vector using T4 DNA ligase (New England Biolabs, Ipswich, USA) with overnight incubation at 16 ^{o}C. The ligation mixtures were transformed

into *E. coli* XL-10 strain and plated on LB agar plates containing 100 μg/ml carbenicillin (Carbenicillin DIRECT, UK).

4.5 Analysis of the constructs

The colonies were screened with colony PCR, which the condition is shown in Table 11. Colonies were grown in 0.5 ml LB media pre-cultures for 3 – 4 hours at 37 ^{o}C. PCR was performed using Dream taq polymerase (Fermentas, Thermo Fisher Scientific Inc., Waltham, USA) with 0.5 μl of pre-cultures as template. Primers used in the reaction were pCoA-Hat_seqF for pCoAT7H3F-hCAR-LBD-hPGC1α constructs and pCoA-3xFlag_seqF for pCoAT7H3F-hPGC1α constructs. T7termR was used for both reactions as 3'-primer. Primers used in the PCR are shown in Table 12. The positive colonies were inoculated and grown in 20 ml of super broth (SB) media containing carbenicillin. The constructs were extracted from the cultures using midiprep DNA extraction kit (Qiagen). Constructs were further analyzed with restriction enzymes selected to cut out the insert using the cloning sites or to digest the construct into well detectable fragments in gel electrophoresis. The correct constructs were confirmed by sequencing at DNA Analysis Facility of Yale University, USA.

Table 11. Colony PCR conditions.

Reagent	Amount	Cycle/Temperature		Time
Culture from a colony	0.5 μl	94 °C		2 min
Dream taq buffer (10 x)	2.0 μl	94 °C		30 s
dNTP (10 mM)	0.4 μl	55 °C	35 x	30 s
Primers (10 μM)	0.4 μl	72 °C		1 – 3 min*
Dream taq polymerase (2U/μl)	0.1 μl	72 °C		5 min
mQH_2O	16.6μl	72 °C		∞
Total volume	20.0 μl			

*** Amplification time was 1/kb min for small PCR products (1000 bp) and 3 min/kb for big PCR products (> 2500 bp).**

Table 12. Primers used in colony PCR and for sequencing.

Name	Sequence (5'-3')
pCoA-Hat_seqF	GAAAGATCATCTGATTC
pCoA-3xFlag_seqF	GGACTACAAAGACCACG
T7termR	GCTAGTTATTGCTCAG

4.6 Mini scale protein expression tests

Constructs were transformed into *E.coli* BL21(DE3)-Gold (Stratagene, Agilent Technologies, Inc., Santa Clara, USA), BL21(DE3)-pRARE and OrigamiB(DE3) (Novagen EMD Millipore, Billerica, USA) strains for mini-scale protein expression tests. The protein expression tests were performed as mini-scale expression tests in a 2 ml cultures. Different expression conditions were tested (Table 13). Pre-cultures were grown overnight at 37 °C 200 rpm and in the morning OD_{600} were measured. The pre-cultures were diluted to OD_{600} 0.1 with media containing half amount of antibiotics, 50 μg/ml carbenicillin and 17 μg/ml chloramphenicol and incubated at 37 °C 200 rpm until OD_{600} reached 0.6 – 0.8. Protein expression was induced with isopropyl-β-D-1-thiogalactopyranoside (IPTG) and the culture was divided and further incubated at 18 and 25 °C overnight at 200 rpm shaking. Cells were harvested by centrifugation at 4300 rpm 30 min and pellets were resuspended in lysis buffer (1 x PBS with 250 mM NaCl, 5 % glycerol, pH 7.0). Lysis buffer volumes were calculated according to the density of the cultures. 40 μl of the cell suspension was pipetted to microcentrifuge tubes where 10 μl of 5 x protein loading buffer (156 mM Tris-HCl pH 6.8, 5 % SDS, 25 % glycerol, 25 % sacharose, 1mM TCEP, 0.1 % bromphenol blue) was added. Cells were lysed at 99 °C for 10 min and centrifuged at 14680 rpm for 2 min. The proteins from the supernatants were separated using sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). To the 12 % or 15 % SDS-PAGE gel 5 – 10 μl of the supernatant was loaded depending on the size of the combs used.

Table 13. Different strains and conditions for mini scale protein expression tests.

Expression conditions	
***E. coli* strains**	BL21(DE3) Gold BL21(DE3) with pRARE OrigamiB(DE3) OrigamiB(DE3) with pLEMO
Incubation time and temperature	2 h at 37 ^{o}C overnight at 25 ^{o}C overnight at 18 ^{o}C 4, 18.5 & 24 h at 30 ^{o}C (only with pLEMO)
Amount of IPTG	titration (0, 0.2, 0.4, 0.6, 0.8 and 1 mM) 1 mM
Growth medium	LB N-Z-amine

OrigamiB(DE3) was tested with and without pLEMO plasmid (Xbrane Bioscience AB, Stockholm, Sweden) for regulating the T7 RNA polymerase by T7 lysozyme induced with rhamnose (Figure 12).

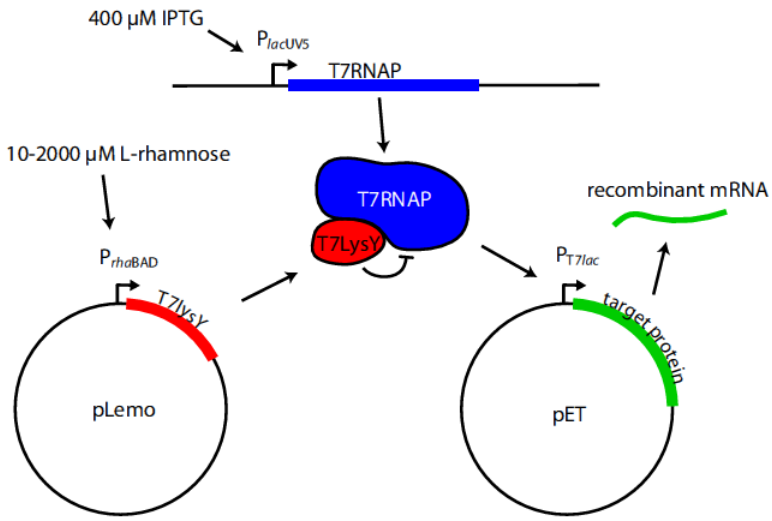


Figure 12. Schematic overview of pLemo system in protein production. Rhamnose is used to induce the T7 lysozyme (T7LysY), which inhibits the T7 RNA polymerase (T7RNAP) and thus regulating the levels of T7RNAP and the expression of the target protein. In the experiment pCoAT7H3F vector was used instead of pET (Wagner *et al.* 2008).

The pLemo system was tested for pCoAT7H3F-CAR-LXXLL, pCoAT7H-CAR-3F-RS and pCoAT7H3F-CAR according to Xbrane protocol (http://www.xbrane.com/files/documents /Protocol_Xbrane_Lemo_System.pdf). The pLemo plasmid was transformed into OrigamiB(DE3) with pCoAT7H3F constructs using heat shock and plated in LB agar plates containing 100 µg/ml of carbenicillin and 34 µg/ml of chloramphenicol to amplify and maintain the plasmids. One colony of each construct was inoculated into 2 ml of N-Z-amine media with antibiotics and grown overnight. The next day 0.2 ml of starter cultures was inoculated to expression 10 ml cultures of N-Z-amine with 50 µg/ml of carbenicillin and 17 µg/ml of chloramphenicol containing 0, 0.1, 0.25, 0.5, 0.75, 1.0 and 2.0 mM of L-rhamnose. In addition a culture without L-rhamnose and without IPTG was included as control. The cultures were further incubated at 30 °C for 4.5 h until OD_{600} reached 0.4 – 0.8. The protein expression was induced with 40 µl of 100 mM IPTG to a final concentration of 0.4 mM IPTG in the cultures. The cultures were then incubated at 30 °C and 3 ml of cultures was collected after 4, 18.5 and 24 h. Cells were harvested by centrifugation 4300 rpm for 30 min at 4 °C

after OD_{600} measurement for the calculation of lysis buffer. The protein samples were prepared as previously described.

4.7 Western blot

Western blotting was used to detect the expressed proteins. Rabbit anti-FLAG antibody conjugated to DyLight™ 800 dye for FLAG fusion proteins and rabbit anti6xHis antibody conjugated to DyLight™ 680 dye for 6xHis conjugated proteins (Rockland Immunochemicals Inc., Gilbertsville, USA) were used for the detection of 3xFLAG and HAT recombinant proteins, respectively.

Western blot was performed from sonicated and centrifuged *E. coli* cultures from the mini scale expression tests. The expression cultures were suspended in lysis buffer and stored at -20 °C. The cultures were sonicated on ice with 25 % amplitude and 2 s pulse on and 2 s pulse off for 1 min using Sonics sonicator tapered four-head probe (Vibra Cell™, Sonics & Materials, Inc., Newtown, USA). From the sonicated lysates 50 µl were taken as crude lysate samples and the rest were centrifuged 20 124 x g (15 000 rpm) for 40 min at 4 °C (Heraus Megafuge 1.0R, Thermo Fisher Scientific Inc., Waltham, USA). Protein concentrations were measured from the crude lysates and supernatants using Bio-Rad protein assay (Bio-Rad Laboratories, Hercules, USA). Measurement was performed with 5 µl of samples and 1 ml of Bio-Rad protein assay reagent. The samples were incubated for 5 min and absorbance was measured at 595 nm using Genesys 10 uv spectrophotometer (Thermo electron corporation, Thermo Fisher Scientific Inc., Waltham, USA). The protein concentrations were calculated using equation from previously measured bovine serum albumin standard curve.

5 and 10 µg of proteins were loaded from the supernatant samples and 2.5 µg from the crude lysate samples to a 10.5 % SDS-PAGE gel that gel was run at 180 V for 1 h. After the gel electrophoresis, proteins were transferred to a PVDF membrane with at 100 V for 50 min and the membranes were blocked using 3 % non-fat milk in PBS overnight at 4 °C. Antibodies were diluted to 3 % milk in PBS to 1:10 000 and anti-FLAG was tested also with 1:15 000 dilution. Membranes were incubated 1 h with antibodies at room temperature and washed with PBS containing 0.1 % Tween-20. Membranes were stored in PBS and scanned using

Odyssey® imager (LI-COR Biosciences, Lincoln, USA) with 700 nm and 800 nm excitation filter.

4.8 Co-purification of hCAR and hPGC1α domains

Protein purification of hCAR and hPGC1α domains was performed from BL21(DE3)-pRARE *E. coli* strain. Purification was tested for hCAR-RS with two different expression conditions such as 2 h at 37 °C and overnight at 20 °C. Proteins were expressed in 10 ml volumes with addition of 1 mM IPTG when OD_{600} reached 0.4-0.6 and 5 ml volumes for negative expression controls without IPTG. Purification was performed using CAR purification buffer including 50 mM Tris-HCl, 100 mM NaCl and 20 % glycerol with pH 8.0 and using PBS buffer including 1x PBS, 250 mM NaCl and 5 % glycerol with pH 7.0. Proteins were purified using Co^{2+} -chelated affinity resin (TALON® Superflow™ Metal Affinity Resin, Clontech, Saint-Germain-en-Laye, Francen). In addition, purification was performed using protease inhibitors in the CAR purification buffer in order to determine whether there was protease activity occurring during the purification process.

Expression cultures were resuspended in purification buffer according to their OD_{600} values and were sonicated for 1 min with pulse on 2 sec and off 2 sec with 25 % amplitude on ice. 40 µl of the sonicated samples were taken as a crude lysate sample for SDS-PAGE analysis. The rest of the sample was centrifuged at 15 000 rpm for 40 min at + 4 °C. Supernatant was collected and 40 µl was taken for SDS-PAGE. The affinity resin was equilibrated by repeated (3x) washing with 10x bead volume of purification buffer. 10 µl of beads were used to purify the proteins from 10 ml cultures. Supernatant was pipetted to the microcentrifuge tube containing the beads and the proteins were allowed to bind to the beads for 1 h at + 4 °C on a platform shaker. The samples were centrifuged at 700 x g for 5 min at + 4 °C and 40 µl of the supernatant was taken as “flow-through” sample for SDS-PAGE indicating the amount of proteins, which did not bind to the beads. Rest of the supernatant was removed from the beads. The beads were washed three times with 10 x beads volume of purification buffer. From the first faction 40 µl was taken for SDS-PAGE. The beads were resuspended with 10 µl of purification buffer and 20 µl of 2x protein loading buffer was added. For the 40 µl SDS-PAGE samples 10 µl of 5x protein loading buffer was added. The samples were heated at 99 °C for 10 min and centrifuged at 14 680 rpm for 5 min. From the crude lysate and bead

samples 5 µl and from the supernatant, flow-through and wash samples 10 µl was loaded to the 12 % SDS-PAGE gel. SDS-PAGE was run at 200 V for 50 min.

4.9 Cell culturing and reporter gene assays

Reporter gene assay was used to investigate the effect of the full-length hPGC1α for hCAR activation. MCF-7 and C3A cell lines were used in the reporter gene assays. Cell cultures were incubated at 37 °C with 5 % CO_2. Transfections were performed using polyethyleneimine (PEI, Polysciences Inc., Warrington, USA) as DNA carrier. The plasmids used in the reporter gene assays are shown in Table 14. The luciferase activities were divided with the protein concentrations measured using Bradford method from every sample in order to decrease the variation caused by the cell amounts between different wells.

Table 14. Plasmids used in the reporter gene assays.

Plasmid	Origin
pGL3-empty	Promega Corporation, Madison, WI, USA
pGL3-PBREM	Dr. Honkakoski (UEF, Kuopio, Finland)
pSG5-hCAR FL	Dr. Honkakoski (UEF, Kuopio, Finland)
pcDNA3-hPGC1α FL	Dr. Billas (IGBMC, Strasbourg, France)

4.9.1 Optimization of reporter gene assay

Different amounts of plasmids and DNA-PEI ratios were tested for optimization of reporter gene analysis. 1:1 and 1:2 DNA-PEI ratios were tested to optimize the transfection. The optimal amount of hCAR was determined by titration with constant amount (500 ng) of PBREM response element using ratios shown in Table 15. pGL3-plasmid without PBREM response element was used as a control alone and with pSG5-hCAR.

Table 15. Titration of CAR with 500 ng of PBREM using MCF-7 and C3A cells.

pSG5-hCAR FL	PBREM-hCAR ratio
50 ng	10:1
100 ng	5:1
250 ng	2:1
500 ng	1:1
1000 ng	1:2

After the amount of hCAR was optimized, titration of hPGC1α was performed. The titration using MCF-7 cells is summarized in Table 16. Two different hCAR concentrations were tested, because both 50 ng and 100 ng in the titration for hCAR gave similar level of high activation.

Table 16. Titration of PGC1α with 50 ng and 100 ng of CAR using MCF-7 cells.

pcDNA-hPGC1α		hPGC1α-hCAR- ratio
50 ng of hCAR	100 ng of hCAR	
10 ng	20 ng	1:5
25 ng	50 ng	1:2
50 ng	100 ng	1:1
100 ng	200 ng	2:1
200 ng	500 ng	5:1
500 ng	1000 ng	10:1

The titration of hPGC1α was performed using C3A cells with 500 ng hCAR and DNA-PEI ratio of 1:1, which is summarized in Table 17. The maximum hPGC1α-hCAR ratio was 2:1 due to the high concentration of total DNA in the cells. Therefore the 5:1 and 10:1 ratios used in the PGC1α titration in MCF-7 cells were not included.

Table 17. Titration of hPGC1α with 500 ng of hCAR in C3A cells.

pcDN-hPGC1α	hPGC1α-hCAR- ratio
50 ng	1:10
100 ng	1:5
250 ng	1:2
500 ng	1:1
1000 ng	2:1

For the reporter gene assays the statistical analysis was performed using GraphPad Prism software (GraphPad Software Inc., La Jolla, USA). Results were analyzed using paired, two-tailed Students' t-test with 95 % of confidence interval. In the hCAR titration experiments the samples containing CAR were compared to the PBREM response element alone and in the hPGC1α titration the samples were compared to PBREM response element with hCAR alone.

5 RESULTS

5.1 Comparison of human and mouse PGC1α homology using BLAST

A comparison for human and mouse PGC1α homology was done in order to determine its conservation of in different species. A sequence alignment was performed using NCBI BLAST tool for nucleotide as well as amino acid sequences. The created alignments are in the Appendix B. The homology in the amino acid sequences for full-length PGC1α orthologues was 94 %, for LXXLL domain 93 % and for RS domain 97 % (Table 21). All three LXXLL motifs, LLAVL, LKKLL and LLKYL, were conserved and also the RS domain, RS motif and RRM motif were unaltered. However, the PPARγ interaction domain contained six substitutions.

Table 21. Comparison of human and mouse PGC1α amino acid sequence homology for the full-length, LXXLL and RS domains.

Sequence	Length	Homology
PGC1α full-length	797 aa	94 %
LXXLL motif	506 aa	93 %
RS domain	291 aa	97 %

5.2 Analysis of the constructs

Colonies were screened using colony PCR with primers amplifying the insert inside the vector. The sizes of PCR product are shown in Table 18. pCoAT7H3F-hPGC1α constructs were amplified using a primer annealing to the FLAG-site and pCoAT7H3F-hCAR-LBD constructs using a primer annealing to the HAT-site. From every construct that produced the correct PCR positive clones were identified.

Table 18. Sizes for the colony PCR products.

Construct		Size of the colony PCR product
#	Name	
1	pCoAT7H3F-CAR	1109 bp
2	pCoAT7H3F-CAR-FL	3505 bp
3	pCoAT7H3F-CAR-LXXLL	2635 bp
4	pCoAT7H3F- CAR-RS	1987 bp
5	pCoAT7H3F-CAR-LXXLLsyn	1566 bp
6	pCoAT7H3F-FL	2550 bp
7	pCoAT7H3F-LXXLL	1669 bp
8	pCoAT7H3F-RS	1013 bp
9	pCoAT7H3F-LXXLLsyn	592 bp

The positive clones from the colony PCR were further analyzed with restriction enzymes. The plasmids were extracted using miniprep columns (Qiagen, Hilden, Germany) and digested using restriction enzymes, which results in fragments detectable in gel electrophoresis and differs from a digestion profile of an empty vector. The correct fragments of the digestions are shown in table 19. The correct digestion profiles were obtained in every construct and from these, depending on the total number of candidate constructs, one or two constructs were sequenced.

Table 19. Sizes of the expected fragments in the restriction analysis of the constructs.

Construct		Restriction enzymes used in plasmid digests	Expected fragments (bp)
#	Name		
1	pCoAT7H3F-CAR	*Bam*HI/*Nde*I *Nco*I	746 & 5396 56, 467 & 5619
2	pCoAT7H3F-CAR-FL	*Bam*HI/*Nde*I *Nco*I	746 & 7784 56, 467, 1518 & 6489
3	pCoAT7H3F-CAR-LXXLL	*Bgl*II	318, 465, 1435 & 5442
4	pCoAT7H3F- CAR-RS	*Bam*HI/*Nde*I *Nco*I	746 & 6266 56, 467 & 6489
5	pCoAT7H3F-CAR-LXXLLsyn	*Bam*HI/*Nde*I *Nco*I	746 & 5840 56, 467 & 6063
6	pCoAT7H3F-FL	*Xba*I *Nco*I/*Xba*I *Hind*III/*Xho*I	1182 & 7348 298, 884, 1220 & 5390 2411 & 6119
7	pCoAT7H3F-LXXLL	*Nco*I *Eco*RI/*Sal*I	1518 & 5404 820 & 6102
8	pCoAT7H3F-RS	*Nco*I/*Xba*I *Hind*III/*Xho*I	876 & 5398 893 & 5381
9	pCoAT7H3F-LXXLLsyn	*Mlu*I/*Xho*I *Nde*I/*Xba*I *Bam*HI/*Xba*I *Bam*HI	1123 & 5463 1414 & 5172 1406 & 5180 6586 (linearization)

The sequencing results confirmed that pCoAT7H3F-hPGC1α-LXXLL, pCoAT7H3F-hPGC1α-RS, pCoAT7H3F-hCAR-LBD, pCoAT7H3F-hCAR-LBD-hPGC1α-LXXLL and pCoAT7H3F-hCAR-LBD-hPGC1α-RS constructs were correct. The pCoAT7H3F-hPGC1α-FL and pCoAT7H3F-hCAR-LBD-hPGC1α-FL did not contain full-length hPGC1α insert.

5.3 Mini scale protein expression tests

Different *E. coli* strains and conditions were tested to optimize the protein expression for hPGC1α-LXXLL, hPGC1α-RS, hPGC1α-LXXLLsyn with and without hCAR-LBD in pCoAT7H3F co-expression vector. In addition, hCAR alone in pCoAT7H3F vector was tested. The protein sizes are shown in Table 20 3xFLAG-linked hRXR cloned in pCoAT7H3F vector was taken as a positive control for expression. Proteins from the expression cultures were separated using denaturing gel electrophoresis, SDS-PAGE.

Table 20. Size of the HAT and 3xFLAG fusion proteins.

Protein	Size (kDa)	Size without tag (kDa)
HAT-CAR LBD	32.4	28.2
3xFLAG-LXXLL	60.2	56.2
3xFLAG-RS	38.9	35.0
3xFLAG-LXXLLsyn	21.0	17.0
3xFLAG-RXR	30.7	26.7

The SDS-PAGE analysis showed that the expression level of the studied proteins was low. Only LXXLLsyn was clearly visible in the gels as a 25 kDa band (Figure 13). Some indications of protein expression were detected with RS and CAR. CAR seemed to be around 25 to 30 kDa and therefore the co-expression of synthetic LXXLL and CAR was difficult to distinguish (Figure 13). The wild-type LXXLL was not detected in any of the expression tests. hRXRα was expressed using the same conditions as for the other proteins.

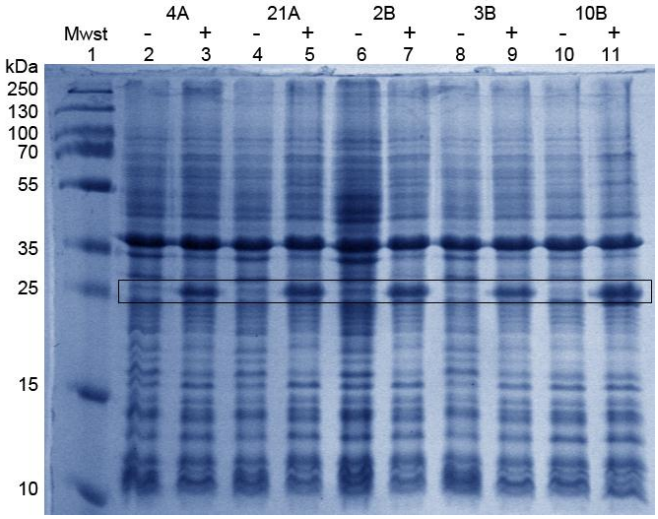


Figure 13. SDS-PAGE of pCoAT7H3F-LXXLLsyn protein expression test for five different colonies. Negative control samples without IPTG are indicated with minus. 15 % gel was used. LXXLLsyn indicated with black rectangle.

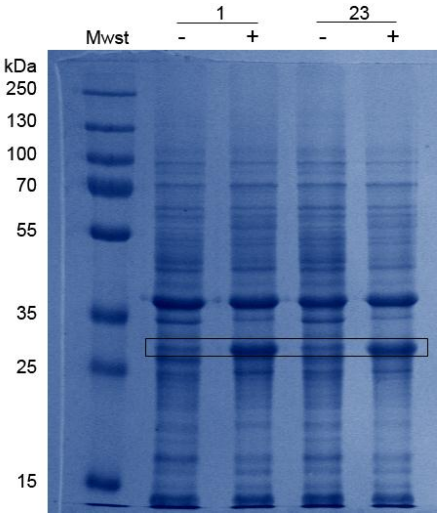


Figure 14. SDS-PAGE of pCoAT7H3F-CAR -LXXLLsyn protein expression test for two different colonies. 12 % gel was used.

5.4 Western blot

Western blot was performed with anti-FLAG and anti-6xHis antibodies in order to detect 3xFLAG-tagged hPGC1α proteins and HAT-tagged hCAR LBD. 3xFLAG-hRXRα and 6xHis-VDR were taken as control samples. 3xFLAG-hRXRα was shown in SDS-PAGE with Coomassie Blue staining and therefore it was included as positive control for anti-FLAG. 6xHis-VDR was taken as a control for anti-6xHis to compare binding of the antibody to HAT and 6xHis.

Anti-6xHis was unable to detect both the HAT-tagged hCAR or the 6xHis-VDR. 3xFLAG fusion proteins were detected from the protein expression test samples (Figure 15 A). 3xFLAG-hRXRα was detected as a clear band and it was around 35 kDa (lane 8). The 3xFLAG-tagged LXXLL region is 60 kDa and there were various bands visible in the 800 nm scan. However, a band around 55-58 kDa visible in LXXLL and faintly in hCAR-LXXLL samples (lanes 2 and 3), which was not present in the other samples. The RS samples contained a band around 40 to 46 kDa, which was consistent with all the samples (lanes 4 and 5).

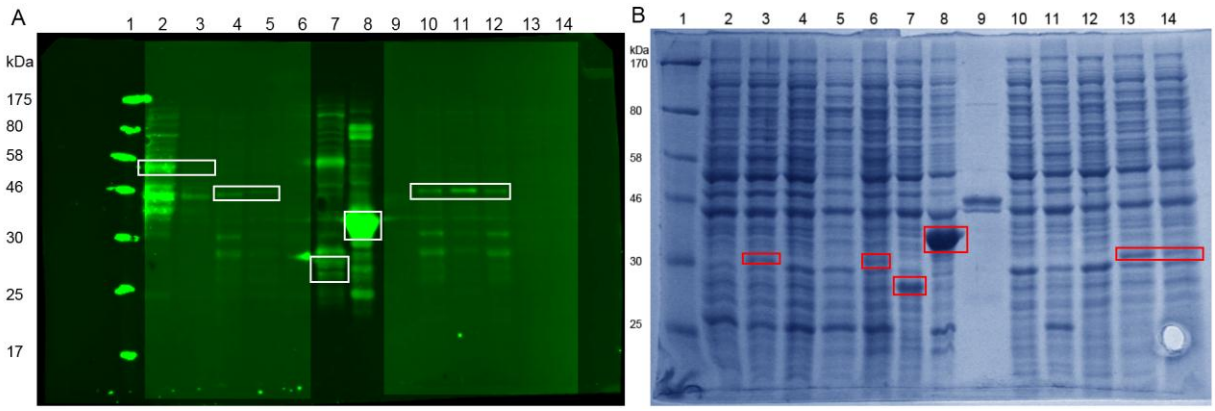


Figure 15. Western blot (A) and Coomassie Blue staining (B) for 10 µg of proteins from supernatant fractions. Western blot was performed with anti-FLAG antibody at a dilution of 1:15000, scanned at 800 nm wavelength. 1) protein marker, 2) LXXLL, 3) CAR-LXXLL, 4) RS 2, 5) CAR-RS, 6) CAR 10, 7) LXXLLsyn, 8) RXR, 9) 6xHis-VDR, 10) RS 3, 11) RS 4, 12) RS 5, 13) CAR 3, 14) CAR 6. Adobe Photoshop was used to enhance the lanes by changing the contrast levels of the image histogram (A, lanes 2-6 and 9-14).

The same supernatant samples used in the western blot were run on 10 % SDS-PAGE with 10 µg of protein and stained with Coomassie Blue (Figure 15 B). Amount of protein in the

hRXRα sample was reduced to 2 µg due to the intense band in western blot. Only hRXRα (lane 8) and LXXLLsyn (lane 10) were clearly visible on the gel using this staining. In addition, the purified 6xHis-hVDR had shown a clear band, which was not detectable in the Western blot membrane and this shows that the anti-6xHis antibody does not seem to work. A weak band of 30 to 32 kDa was visible in hCAR-LXXLL sample (lane 3), in CAR-RS sample (lane 6) and in CAR samples (lanes 13 and 14), which seems to be hCAR. LXXLL and RS could not be clearly distinguished from the other *E. coli* proteins.

Due to the weak detection of the recombinant proteins in the supernatant samples, western blot was repeated using crude lysate samples. Western blot was performed with 2.5 µg of protein and Coomassie Blue staining was performed with 5 µg of protein for the same samples (Figure 16). In the hRXRα sample, only 1 µg of protein was taken for western blot and 2 µg for Coomassie Blue staining. There was a 55-58 kDa band visible in LXXLL sample and very weakly in CAR-LXXLL sample (Figure 16 A, lane 3 and 4), similar to supernatant samples. RS was clearly visible in all RS and CAR-RS samples with 40 to 46 sized band on the Western blot membrane and Coomassie gel (Figure 16 A and B, lanes 5 – 10). CAR was visible in CAR-RS samples around 28 to 30 kDa with Coomassie Blue staining (B, lanes 9 and 10). LXXLLsyn was weakly detected around 26 to 28 kDa with Western blot and Coomassie Blue staining (A and B, lane 11). hRXRα was clearly visible around 32 to 35 kDa with western blot and Coomassie staining (A and B, lane 12).

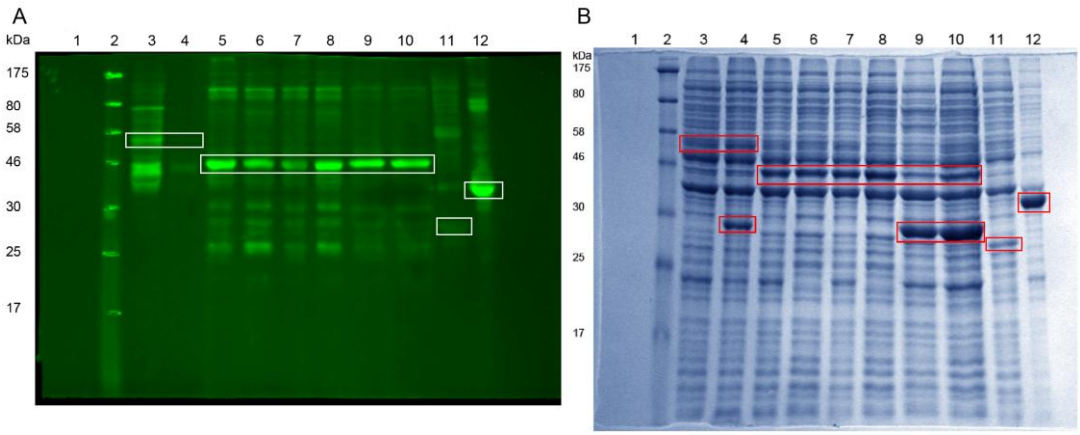


Figure 16. Western blot and Coomassie Blue staining for crude lysate fractions. A) Western blot was performed with 2.5 µg of proteins with anti-FLAG antibody at a dilution of 1:25000 scanned at 800 nm wavelength and B) Coomassie staining was performed for 5 µg of proteins. Protein were separated with 10 % SDS-PAGE. 1) Empty, 2) protein marker, 3) LXXLL, 4) CAR-LXXLL, 5) RS 2, 6) RS 3, 7) RS 4, 8) RS 5, 9) CAR-RS A, 10) CAR-RS B, 11) LXXLLsyn, 12) RXR 1 µg in western blot and 2 µg in Coomassie staining.

5.5 Co-purification of hCAR and hPGC1α domains

Purifications of hCAR-hPGC1α-RS, hCAR-hPGC1α-LXXLL and hCAR-hPGC1α-LXXLLsyn complexes and hCAR alone were performed on lysates derived from 10 ml cultures using metal affinity beads binding to the HAT fused to hCAR. Expression was performed for 2 h at 37 ºC or overnight at 20 ºC. hCAR-hPGC1α-RS purification was performed using CAR purification buffer with pH of 8.0 for both expression conditions (Figure 17). The theoretical size for CAR with HAT tag is around 32 kDa and CAR was purified as a clear fraction around 25 to 30 kDa and. However, there were three bands scattered around that size and the bands were intensive in the purified fraction (lanes 7 and 13). A band of around 40 kDa was detected in crude lysate and bound protein samples with 20 ºC overnight expression (lanes 9 and 13). This protein band may represent hPGC1α-RS, which has a theoretical size of 39 kDa. However, this band was absent from the 2 h at 37 ºC expression samples. In addition, purification for the same expression cultures was performed using PBS buffer with pH of 7.0 in order to determine the effect of the buffer to the purification. There were no clear advantages for using PBS compared to CAR purification buffer (data not shown) and therefore for the rest of the purifications CAR purification buffer was used. Expression level of the desired proteins seemed to be higher at 20 ºC with overnight incubation compared to 2 h at 37 ºC expression condition.

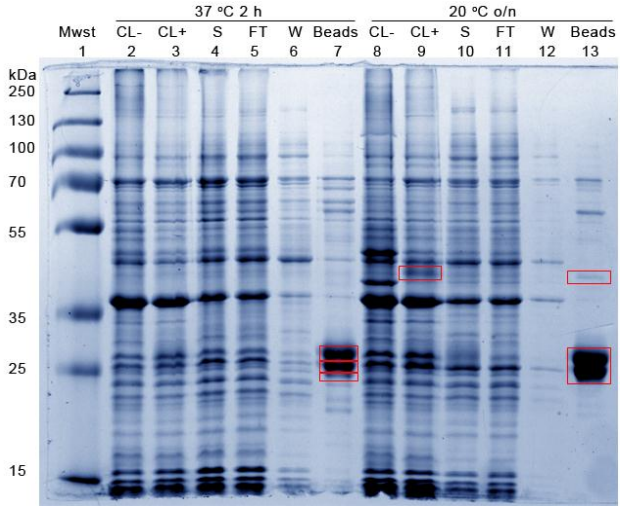


Figure 17. Purification of CAR-RS protein complex using CAR purification buffer. Cultures without IPTG were included as negative expression controls, indicated with "–" in crude lysate samples. Other purification steps were performed only for positive expression samples. CL = crude lysate, S = supernatant, FT = flow-through, W = wash, Beads = bound protein.

Due to the multiple bands in the 25 to 30 kDa range, purification using protease inhibitors was tested to determine whether the bands could be proteolysed fragments of hCAR. The purification was performed in the presence and absence of hPGC1α-LXXLLsyn domain. However, the three bands were also detected in the presence of protease inhibitors and therefore they may be formed during the expression in *E. coli* (Figure 18). Comparing the purifications of hCAR in the presence and absence of hPGC1α-LXXLL, there seems to be only small difference. The expression in hCAR-LXXLL sample is lower and there are no additional bands indicating the expression of the LXXLL domain. However, in the purification samples of LXXLLsyn with hCAR, the expression level is higher than with the other co-expressed proteins. Only with CAR-LXXLLsyn sample the expression can be detected from the crude lysate samples with IPTG (Figure 18, lane 9). Moreover, there is a strong band around 55 to 60 kDa, which is not visible with the other proteins. It appeared to be a double band, however the lower band is hardly detectable.

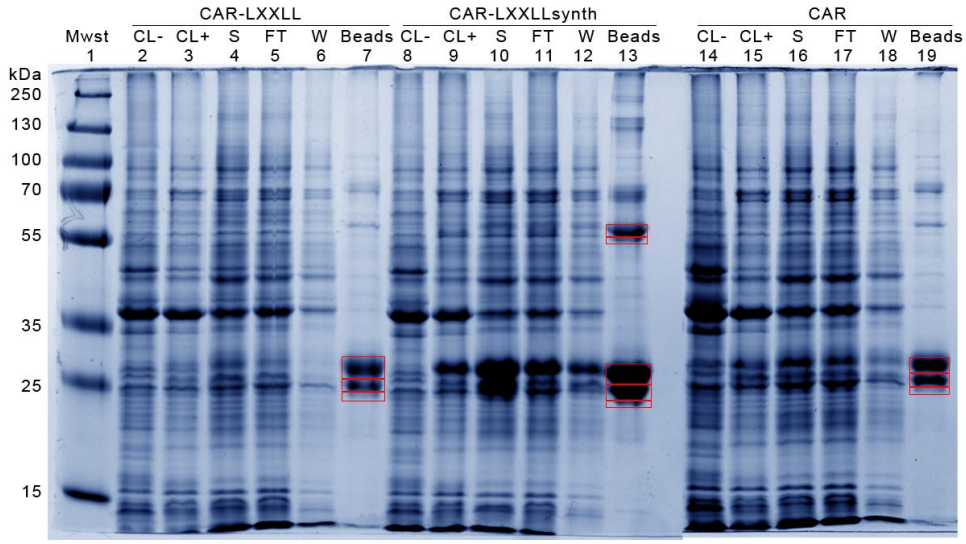


Figure 18. Purification of CAR-PGC1α-LXXLL, CAR-PGC1α-LXXLLsyn and CAR. Purification was performed from 10 ml cultures at 20 °C with overnight incubation using CAR purification buffer with protease inhibitors.

5.6 Cell culturing and reporter gene assays

Cell culturing was performed using MCF-7 and C3A cell lines for the reporter gene assays. Microscopic images were taken to visualize the morphology of the two different cell lines (Figure 19). Both are epithelial cells and grow adherent, attached to the culturing dishes. C3A cells are more round-shaped compared to MCF-7 cells.

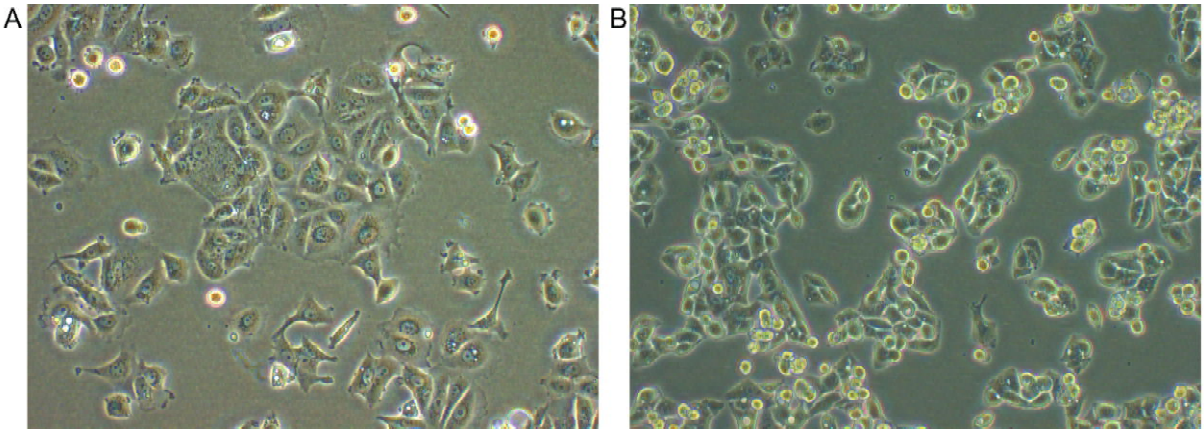


Figure 19. Microscopic images of cell lines used in the assays. A) MCF-7 cells and B) C3A cells taken with 100-fold magnification.

In the reporter gene assays DNA-PEI ratios of 1:1 and 1:2 were tested and the optimal amount of hCAR and hPGC1α plasmids were selected by titration. First hCAR titration was performed and then hPGC1α with selected optimal hCAR amount using MCF-7 cell line. Empty pGL3 vector with and without hCAR and pGL3-tk-luc-PBREM plasmids were included as controls. The results were normalized to protein amounts measured using Bradford method. The result showed that empty pGL3 plasmid did not yield luciferase activity with or without CAR with normalized activity of around 100 units in both DNA-PEI ratios of 1:1 and 1:2 (Figure 20). The luciferase activity of response element with *luciferase* gene, tk-luc-PBREM, alone yielded 2-fold and 1.2-fold increase compared to the samples without the *luciferase* gene with DNA-PEI 1:1 ratio and 1:2 ratio, respectively. The activities with 1:1 ratio were around 3- to 4.5-fold higher compared to 1:2 ratio in the samples containing hCAR. Addition of 50 ng of CAR increased the activity by 5-fold compared to PBREM alone in the DNA-PEI 1:1 samples. The activity of hCAR declined in a dose dependent manner with 1000 ng of hCAR the luciferase activity being only 2-fold compared to PBREM alone. Similar trend was detectable with DNA-PEI 1:2 samples.

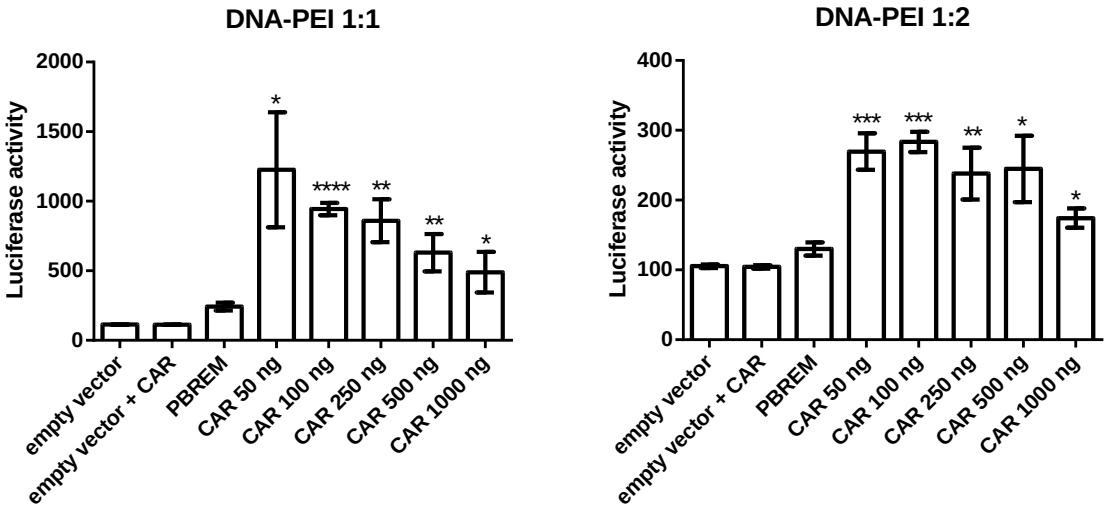


Figure 20. Luciferase activity for CAR titration using MCF-7 cells with mean and standard deviation (n = 3). Titration was performed with DNA-PEI ratios 1:1 and 1:2 and with 500 ng of PBREM or empty vector. Titration of CAR was performed once using triplicates. Samples containing CAR were compared to PBREM alone with t test. * P<0.05, **P<0.01, ***P<0.001, ****P<0.0001

The titration of hPGC1α was performed using 50 and 100 ng of hCAR, in order to verify the results showing higher activity of 50 ng of hCAR. DNA-PEI ratio of 1:1 was chosen for these transfections. With 50 ng of hCAR the activity was 1.5-fold compared to PBREM alone (Figure 21). The activity with 100 ng of CAR was 3-fold compared to PBREM alone. The activity peaked at 2.7-fold with 200 ng of hPGC1α in assay with 50 ng of hCAR. With 100 ng of hCAR, hPGC1α increased the activity to the maximum by only approximately 2-fold using 100 ng of hPGC1α. The further increase of hPGC1α seemed to decrease the activity in a dose dependent manner but with 50 ng of hCAR the amounts of hPGC1α did not seem to affect the activity levels profoundly.

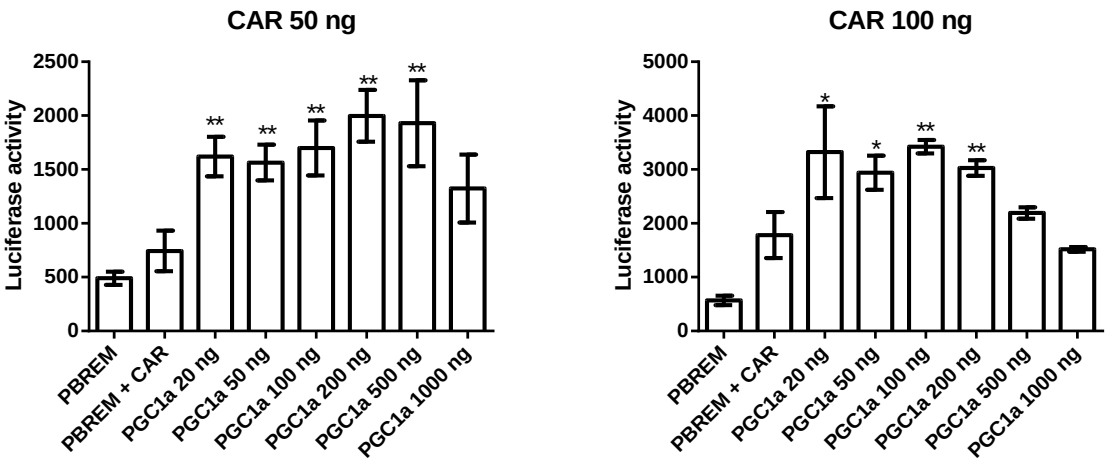


Figure 21. Luciferase activity for PGC1α titration using MCF-7 cells with mean and standard deviation (n = 3). Titration was performed with 50 ng and 100 ng of CAR with 500 ng of PBREM using DNA-PEI ratios 1:1. Experiment was performed twice, data from one experiment is shown. Samples containing PGC1α were compared to PBREM + CAR with t test. * P<0.05, **P<0.01

To compare the results in different cell lines, the reporter gene assays for hCAR titration and hPGC1α titration were also performed in C3A. The hCAR titration was performed as previously described for MCF-7. The DNA-PEI ratio of 1:1 resulted also in higher activities compared to 1:2 ratio, as seen in MCF-7 cells. The highest measured activity was 3.5-fold. On the contrary to best condition where 50 ng of hCAR was used in MCF-7 cells, 500 ng of hCAR was required to yield the highest luciferase activity with 5-fold increase compared to PBREM alone. The empty pGL4 vector in the presence and in the absence of hCAR yielded approximately the same level of luciferase activity.

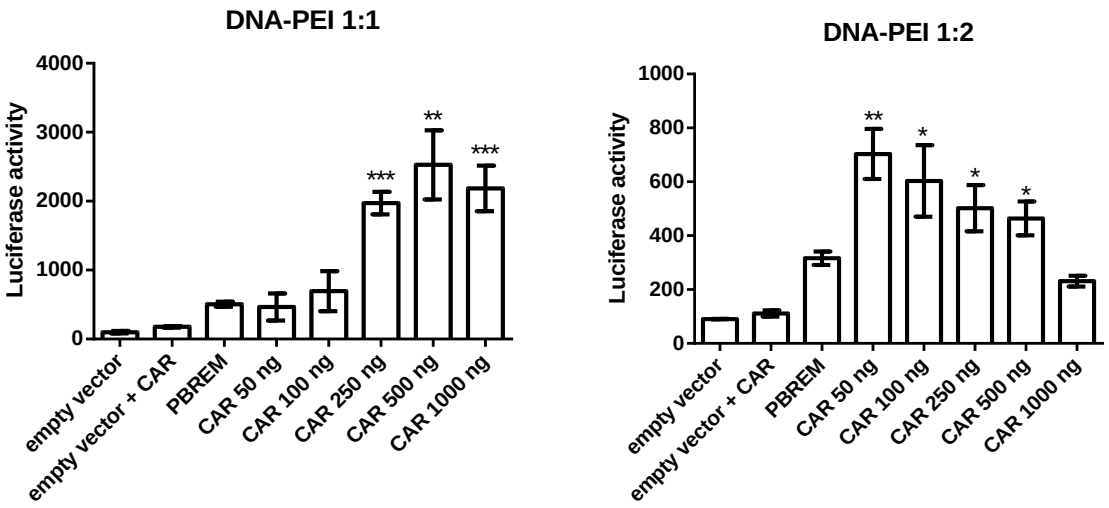


Figure 22. Luciferase activity using for CAR titration C3A cells with mean and standard deviation (n = 3). Titration was performed with DNA-PEI ratios 1:1 and 1:2 and with 500 ng of PBREM or empty vector. Titration of CAR was performed twice times with triplicates, data from one experiment is shown. Samples containing CAR were compared to PBREM alone with t test. * P<0.05, **P<0.01, ***P<0.001

The titration of hPGC1α in C3A cells was performed with 500 ng of hCAR using DNA-PEI ratio of 1:1. Two different amounts of PBREM were tested in order to determine whether 500 ng was optimal for the assays. First, the titration using 500 ng of PBREM was performed as in the previous assays and then the assay was repeated using 100 ng of PBREM (Figure 23). In the assay with 500 ng of PBREM, the activity of hCAR was only 2-fold compared PBREM alone. PBREM resulted in higher activity compared to the previous assays (Figures 21 and 22). The luciferase activity increased only 1.5-fold to its maximum with 100 ng and 1000 ng of hPGC1α compared to hCAR with PBREM. With 100 ng of PBREM, the activity of hCAR was 2.5-fold compared to PBREM alone. Moreover, hPGC1α increased the activity by 5.7-fold with 1000 ng and 3.8-fold with 500 ng compared to hCAR with PBREM. Untransfected cells were taken as a control to determine the background signal of the assay. The activity of cells alone was around 60 absolute light units, which was at the same level as in the samples with 100 ng of PBREM alone, thus in C3A no detectable activity that may come from cells was observed.

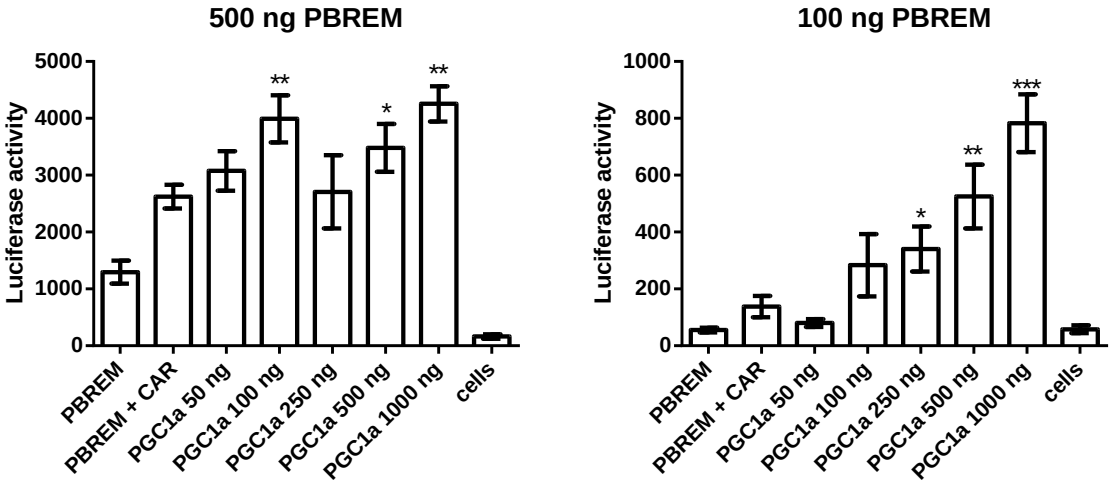


Figure 23. Normalized luciferase activity for PGC1α titration using C3A cells with mean and standard deviation (n = 3). Titration was performed using 500 ng of CAR with 500 ng and 100 ng of PBREM. DNA-PEI ratio was 1:1. Experiment was performed three times with triplicates using 500 ng of PBREM and once using 100 ng of PBREM, data from one experiment is shown. Samples containing PGC1α were compared to PBREM + CAR with t test. * P<0.05, **P<0.01, ***P<0.001

6 DISCUSSION

6.1 Comparison of human and mouse PGC1α homology using BLAST

The comparison of the homology between human and mouse PGC1α orthologues was performed in order to determine the possible differences. Shiraki and colleagues (Shiraki *et al.* 2003) performed their experiments using mouse PGC1α and in this thesis hPGC1α was used. There may be differences in the responses with different species, for example human and mouse CAR respond in different ways to some known hCAR ligands.

The results showed that the similarity between human and mouse full-length PGC1α is 94 % LXXLL motif 93 % and RS domain 97 %. The RS domain seems to be well conserved due to the high similarity. However, the three LXXLL motifs are all unchanged and also the RS domain, but the PPARγ interaction motif contains six substitutions. The alterations between the two PGC1α sequences were for the most part in the regions, which do not contain

important interaction surfaces or functional properties. Therefore it could be presumed that the difference between human and mouse PGC1α is not influential for the interaction studies.

6.2 Molecular cloning of the recombinant protein constructs

hCAR LBD and LXXLL and RS domains of human PGC1α were cloned into co-expression vector using classical cloning. Due to intrinsic *Nco*I site the cloning of the full-length hPGC1α was attempted with the subsequent insertion of both, LXXLL-*Nco*I and RS-*Bam*HI/*Nco*I, into pCoAT7H3F-vector digested with *Nco*I and *Bgl*II. Except this the pCoAT7H3F-hPGC1α full-length construct, all other designed constructs were successfully cloned. Unfortunately, there is no data indicating the reason for this failure. The additional future approaches to obtain the full-length hPGC1α expression construct may be in using ligation independent cloning techniques. However, the current wild-type cDNA fragments that generated the LXXLL and RS domain construct show problems in expression level and protein solubility that is to be expected also for full-length hPGC1α due to the large size and the lack of chaperones in the bacterial expression system.

6.3 Analysis of the constructs

The constructs were analyzed using colony PCR, restriction enzymes and sequencing. Colony PCR is an efficient, fast and relatively cheap method for screening a large number of colonies. A vector-specific T7 terminator and an insert-specific 3xFLAG or HAT primers were used in the PCR to amplify the DNA containing the inserts. Colony PCR identified the correct sized product for pCoAT7H3F-LXXLLsyn and pCoAT7H3F-RS constructs with all of the clones analysed, other constructs produced various products and only few of the clones produced the correct PCR product. For the constructs, where colony PCR produced only correct product with all of the clones, it may be that the ligation was more efficient than for the constructs where only few clones produced the correct product. The regions of the vector, where unspecific primer annealing occurred may have been also could be amplified. In this case, the colony did not contain the correct insert and the PCR reaction may have been shifted towards unspecific amplification.

In further analysis, restriction endonucleases were used to verify the constructs and the size of the inserts. The restriction enzymes were selected either to cut out the insert using the cloning

sites or to digest inside the insert and insert flanking region of the vector in order to create fragments that are well detectable in the gel electrophoresis. The correct digestion profiles were obtained from every construct except the LXXLLsyn. However, there seemed to be a problem with *Xba*I digestion, because it failed to cut any of the LXXLLsyn constructs or the empty pCoAT7H3F vector. Since by codon-optimization procedure all the intrinsic sites were removed from this insert, it was difficult to find restriction enzymes adequate for the successful analysis. Despite the lack of the restriction analysis, hCAR LBD was sub-cloned into the pCoAT7H3F-LXXLLsyn construct. The colony PCR and restriction analysis for the CAR-LXXLLsyn construct showed that the construct was correct.

Sequencing was performed with the 3xFLAG or HAT primers from 5' prime and with T7 terminator primer from 3' prime. The T7 primer was added by the sequencing facility. However, none of the T7 terminator sequencing samples resulted any readable sequence. The 3xFLAG and HAT primers were working and the sequencing results confirmed that every construct except the full-length hPGC1α alone or in combination with hCAR were correct. Although the pCoAT7H3F-CAR-PGC1α-FL contained hCAR but the full-length hPGC1α was absent.

6.4 Mini scale protein expression tests

The mini scale expression was analyzed by SDS-PAGE with Coomassie Blue staining for RS and LXXLLsyn domains and possibly hCAR. The expression levels of the proteins in *E. coli* were small for all proteins except the LXXLLsyn domain. In addition, 3xFLAG-hRXRα in the same co-expression vector was taken as a positive control for the expression. The expression of LXXLLsyn and hRXRα showed that the 3xFLAG fusion in the co-expression vector was functional. However, there was no clear co-expression of hCAR with any of the 3xFLAG-tagged proteins. This could be due to that the *E. coli* could not process the expression from two T7 initiator sites or the equilibrium was shifted to expressing only the one of the two proteins.

The expression level for the LXXLLsyn was higher compared to LXXLL or RS domains. This could indicate that the codon optimization for the synthesized LXXLL gene aided the expression process in *E. coli.* However, the wild-type LXXLL and LXXLLsyn do not have the same size thus the direct comparison of the expression between these two constructs is

difficult. hPGC1α contains rare codons for *E. coli* and this could decrease the protein expression level. A rare codon analysis for 3xFLAG-LXXLL and 3xFLAG-RS was performed using online tool at Genscript website (http://www.genscript.com/cgi-bin/tools/rare_codon_analysis). The results showed that both recombinant proteins are not optimal for expression in *E. coli* host. The Codon Adaption Index (CAI) indicates the frequency of codons used in translation from mRNA to protein with CAI of 1.0 being the ideal expression level and >0.8 is considered as good for expression. 3xFLAG-LXXLL resulted in CAI of 0.59 and 3xFLAG-RS in 0.61. The protein expression was tested with different *E. coli* strains and one was BL21(DE3) with pRARE plasmid. pRARE contains genes for tRNAs with codons rare for *E. coli* and therefore should enhance protein expression (Novy *et al.* 2001). In addition, the prediction of protein disorder for hPGC1α using DISOPRED version 2 server (http://bioinf.cs.ucl.ac.uk/disopred/) indicated that in most of the protein sequence hPGC1α is an intrinsically disordered protein This further hinder its expression in E. coli and the sub-optimal solution is to use folded domain parts with co-expression of the partner proteins. Moreover, there is a possibility to shift to different host such as yeast or insect cells that may provide better result in terms of protein quality and yield.

The calculated protein size for HAT-hCAR is approximately 32 kDa. In the protein expression test, hCAR was visible around 30 kDa in the 12 % SDS-PAGE. The calculated protein size for 3xFLAG-LXXLLsyn is approximately 21 kDa, but it seemed to migrate around 25 kDa. The small difference between the sizes of the two proteins caused difficulties in their detection from the hCAR and LXXLLsyn co-expression tests. The hCAR-LXXLLsyn co-expression sample was separated with 16 % SDS-PAGE and ran for 75 and 105 min in order to detect the two proteins (data not shown). However, there seemed to be only one band around 25 kDa. Therefore it can be suspected that there is only LXXLLsyn expressed.

6.5 Western blot

The Coomassie Blue staining was possibly not sensitive enough to detect the low levels of protein expression and therefore Western blot was performed. Western blotting was also performed in order to verify the protein sizes of 3xFLAG- and HAT-tagged proteins using antibodies against FLAG and 6xHis. HAT tag is bigger and largely differs from the 6xHis tag and therefore it was expected that the antibody may not work for the HAT fusion proteins. As

expected, there were no visible bands after the 700 nm scan. However, the 6xHis antibody seemed not to be able to detect the 6xHis-VDR in any of the experiments that has been visualized with Coomassie Blue staining. However, there is no certainty of why 6xHis-VDR was not visible and better positive controls are needed in future experiments. Western blot was performed for crude lysate and supernatant samples.

The FLAG antibody seemed to bind unspecifically to various proteins blotted to the membranes and therefore the detection of the 3xFLAG-tagged LXXLL and RS domains was difficult. The 3xFLAG-hRXRα was clearly visible in both crude lysate and supernatant samples due to its high expression, which was also visible with Coomassie Blue staining. In addition, the LXXLLsyn was visible with Coomassie Blue staining and could therefore be distinguished. The detection of hRXRα and LXXLLsyn showed that the Western blot was successful and the FLAG antibody was binding to the 3xFLAG fusion proteins, despite the unspecific binding. The unspecific binding could have been reduced by testing different blocking solutions, for example the Odyssey blocking buffer recommended by the LI-COR company, or using more pure protein samples.

Using the crude lysate samples, RS was detected with western blot and Coomassie Blue staining and hCAR from hCAR-RS was also visible with Coomassie staining. LXXLL was difficult to detect from Western blot membrane due to many unspecific bands. Moreover, LXXLL was not visible with Coomassie Blue staining. However, there seemed to be a band around 60 kDa in the LXXLL and hCAR-LXXLL samples that was detected from the Western blot membrane using crude lysate and supernatant samples. Therefore, there is an indication of protein expression of every recombinant protein used in this study.

6.6 Co-purification of hCAR and hPGC1α domains

Protein purification was performed using metal affinity beads binding to the HAT tag. The main aim was to co-purify hPGC1α domain proteins with HAT-tagged hCAR. This would indicate the interactions between hCAR and hPGC1α LXXL and RS domains.

The results from protein purifications showed that hCAR was purified from all of the samples. However, LXXLL, RS and LXXLLsyn were not clearly visible from the purified

samples. This could be due to the very low level of protein expression and thus not enough LXXLL or RS to interact with hCAR. However, in the co-purification with LXXLLsyn an approximately 55 kDa band appeared which was not visible in other protein purifications. One option could be that the transcription of the mRNA did not stop after hCAR and continued to the end of LXXLLsyn. This would result in a protein of approximately 55 kDa. However, this is highly unlikely due to the presence of astop-codon at the end of hCAR and there a hCAR with the correct size expressed at high level.

Performing the co-purification of hCAR and hPGC1α domains in larger culture volumes could have yielded more hPGC1α and the interaction could have been detected better. There was a band around 40 kDa in co-purification of RS with hCAR (Figure 17, lane 13), which was not clearly visible in the co-purifications for other proteins. The calculated size for RS domain is 39 kDa and therefore it could be that the RS domain was co-purified with hCAR. Shiraki and colleagues (Shiraki *et al.* 2003) showed that also RS domain has the ability to bind directly to hCAR LBD, despite the fact that LXXLL motif is important region for the activation of NRs (Savkur & Burris 2004). The hCAR-LXXLL co-expression plasmid yielded only low level of proteins and this could affect that the binding of LXXLL to hCAR was not detected. However, the LXXLLsyn was expressed in higher level, but due to the close proximity of the LXXLLsyn and the hCAR LBD, the LXXLLsyn was overlaid by hCAR.

6.7 Cell culturing and reporter gene assays

Cell culturing was performed with MCF-7 cell line, because it is a common, easy to use cell line that is widely used for reporter gene assays. Titration of hCAR with pGL3-PBREM using two different DNA-PEI ratios, 1:1 and 1:2, showed that 1:1 ratio and lower hCAR amounts resulted in higher luciferase activity. It has been observed that cell-type specific high DNA doses may create atypical signaling in cells such as quenching of cofactors leading to artifacts. This could be one of the reasons, why high DNA levels of hCAR would not produce higher activities. Therefore 50 and 100 ng of hCAR were tested in further hPGC1α titration.

The results showed that hPGC1α failed to further increase the activity of hCAR and eventually a dose dependent decrease in the hCAR's activity was observed compared to samples with absent hPGC1α. This could be due to the fact that MCF-7 cells contain high

levels of intrinsic coactivators and this may have interfered with hCAR-hPGC1α interaction. For example signaling pathways of the p160 coactivator family have been shown to have a key role in breast cancer cells (Wertheimer *et al.* 2012). Moreover, RAC3 coactivator has been shown to interact with hCAR in a ligand-independent manner (Molnár *et al.* 2005). This could be one factor interfering with the hPGC1α mediated activation of hCAR. Therefore C3A cell line was also tested. C3A cells are hepatocytes and therefore ideal to use for hCAR studies, because hCAR is primarily expressed in the liver *in vivo*.

Results from the reporter gene assays using C3A cells differed from using the MCF-7 cells. The titration of hCAR showed that much higher amount, 500 ng, seemed to be the optimal and DNA-PEI ratio of 1:1 yielded higher activities that 1:2, which was same as with MCF-7. The titration of hPGC1α resulted in only 1.5-fold increase in the activation of hCAR and no dose dependent response was detected. The different amounts of hPGC1α did not shown any specific trends. The highest activities were in samples with 100 and 1000 ng and the lowest activities with 50 and 250 ng of PGC1α.

However, the response element, PBREM, was not optimized and this could have affected the results. Therefore, the titration of hPGC1α using C3A cells was repeated with 100 ng of PBREM. Interestingly, the profile of the titration changed and dose response dependence was detected. The increase of PGC1α increased the activity and 1000 ng of PGC1α resulted in the highest activity with 5.7-fold increase compared to PBREM with 500 ng of hCAR. It seems that the amount of the total DNA and the ratio for different plasmids are important in order to obtain reliable results. It would be beneficial to optimize the amount of every plasmid used in the assay and use the optimal ratio of DNA and transformation reagent.

7 CONCLUSIONS AND FUTURE DIRECTIONS

The aim of this thesis was to investigate the protein-protein interaction of recombinant nuclear receptor hCAR and coactivator hPGC1α using the prepared tools. The large part of this project was the preparation of various expression and co-expression constructs to produce recombinant hCAR and hPGC1α. The reporter gene assays using full-length hCAR and hPGC1α showed that hPGC1α activates hCAR with maximum of 5.7-fold. The next step will be the testing of different ligands in order to determine the possibility of the ligand-dependent

mediation of the protein-protein interaction. The interactions may be investigated using mammalian two-hybrid assay with just LBD domain of hCAR and different fragments of hPGC1α. In a longer time scale this may result in finding novel therapeutical compounds for treatment of various diseases such as diabetes or cardiovascular disease.

The detailed investigation for the impact of hCAR-hPGC1α interaction of different regions of hPGC1α was incomplete due to lack of time. However, there was some indication for interaction of the RS domain with hCAR LBD in the co-purification assays. The expressions and purifications will be done using larger cultures in order to gain more proteins. In addition, the purification will be done for the constructs with hPGC1α domains in the absence of hCAR and comparing these purification profiles with the co-purification profiles of LXXLL and RS domains in the presence of hCAR could high-light the hPGC1α domains in the SDS-PAGE analysis. The interesting triple band around 25 to 30 kDa in the purifications of hCAR without and with hPGC1α proteins will be studied using mass spectrometry. This may reveal whether the bands are fragments of hCAR. Mass spectrometry will be used to investigate the 55 kDa band in hCAR-LXXLLsyn purification, whether it contains the LXXLLsyn or hCAR.

The proteins will be purified in large quantities and will be used in interaction studies for example limited protease digestion, which may give more specific information about the stability of the protein complex. In addition, AlphaScreen will be used to investigate the binding affinities between hCAR and full-length or different motifs of PGC1α. These assays will be performed in the presence of ligands, which may show their effect on the stability of the interaction.

This study describes the preparation of the tools and optimization of the methods needed for the protein-protein interaction studies using hCAR and hPGC1α. Preliminary results show that there is interaction between these two proteins that involves both LXXLL and RS domains. The aim of the further studies may be to identify the specific regions important for this interaction.

8 REFERENCES

Aranda A & Pascual A (2001) Nuclear hormone receptors and gene expression. *Physiol. Rev.* **81**:1269-1304

Baes M, Gulick T, Choi H, Martinoli M, Simha D, & Moore DD (1994) A new orphan member of the nuclear hormone receptor superfamily that interacts with a subset of retinoic acid response elements. *Mol. Cell. Biol.* **14**:1544-1552

Bhalla S, Ozalp C, Fang S, Xiang L & Kemper JK (2004) Ligand-activated pregnane X receptor interferes with HNF-4 signaling by targeting a common coactivator PGC1α. Functional implications in hepatic cholesterol and glucose metabolism. *J. Biol. Chem.* **279**:45139-45147

Blumberg B, Sabbagh W Jr, Juguilon H, Bolado J Jr, van Meter CM, Ong ES & Evans RM (1998) SXR, a novel steroid and xenobiotic-sensing nuclear receptor. *Cell.* **92**:3195-3205

Boström P, Wu J, Jedrychowski MP, Ye L, Lo JC, Rasbach KA, Boström EA, Choi, Lomg Jz, Kajimura S, Zingaretti MC, Vind BF, Tu H, Cinti S, Højlund K, Gygi SP & Spiegelman BM (2012) A PGC1-α-dependent myokine that drives brown-fat-like development of white fat and thermogenesis. *Nature.* **481**:463-468

Chang C, Norris JD, Grøn H, Paige LA, Hamilton PT, Kenan DJ, Fowlkes D & McDonnell DP (1999) Dissection of the LXXLL nuclear receptor-coactivator interaction motif using combinatorial peptide libraries: discovery of peptide antagonists of estrogen receptor α and β. *Mol. Cell. Biol.* **19**:8226-8239

Chen D, Ma H, Hong H, Koh SS, Huang SM, Schurter BT, Aswad DW & Stallcup MR (1999) Regulation of transcription by a protein methyltransferase. *Science.* **248**:2174-2177

Christian M, White R & Parker MG (2006) Metabolic regulation by nuclear receptor corepressor RIP140. *Trends Endocrinol. Metab.* **17**:243-250

Choi, H-S, Chung M, Tzameli I, Simha D, Lee Y-K, Seol W, & Moore DD (1997) Differential transactivation by two isoforms of the orphan nuclear hormone receptor CAR. *J. Biol. Chem.* **272**:23565-23571

Cotnoir-White D, Lapperière D & Mader S (2011) Evolution of the repertoire of nuclear receptor binding sites in genomes. *Mol. Cell. Endocrinol.* **344**:76-82

Desvergne B, Michalik L & Wahli W (2006) Transcriptional regulation of metabolism. *Physiol. Rev.* **86**:465-514

Ding X, Lichti K, Kim I, Gonzalez FJ & Staudinger JL (2006) Regulation of constitutive androstane receptor and its target genes by fasting, cAMP, hepatocyte nuclear factor α, and the coactivator peroxisome proliferator-activated receptor γ coactivator-1α. *J. Biol. Chem.* **281**:26540-26551

Dong B, Saha PK, Huang W, Chen W, Abu-Elheiga LA, Wakil SJ, Stevens RD, Ilkayeva O, Newgard CB, Chan L & Moore DD (2009) Activation of nuclear receptor CAR ameliorates diabetes and fatty liver disease. *Proc. Natl. Acad Sci.U.S.A.*. **106**:18831-18836

Dong B, Qatanani M & Moore DD (2009) Constitutive androstane receptor mediates the induction of drug metabolism in mouse models of type 1 diabetes. *Hepatology.* **50**:622-629

Dunlop T, Väisänen S, Frank C & Carlberg C (2004) The genes of the coactivator TIF2 and the corepressor SMRT are primary 1α,25$(OH)_2D_3$ targets. *J. Steroid Biochem. Mol. Biol.* **89-80**:257-260

Eloranta JJ & Kullak-Ublick GA (2004) Coordinate transcriptional regulation of bile acid homeostasis and drug metabolism. *Arch. Biochem. Biophys.* **433**:397-412

Farnham P (2009) Insight from genomic profiling of transcription factors. *Nat. Rev. Genet.* **10**:605-616

Finck BN & Kelly DP (2006) PGC1 coactivators: inducible regulators of energy metabolism in health and disease. *J. Clin. Invest.* **116**:615-622

Flaus A & Owen-Hughes T (2011) Mechanisms for ATP-dependent chromatin remodeling: the means to the end. *FEBS J.* **278**:3579-3595

Forman BM, Tzameli I, Choi HS, Chen J, Simhak D, Seol W, Evans RM & Moore DD (1998) Androstane metabolites bind to and deactivate the nuclear receptor CAR-β. *Nature.* **395**:612-615

Fowler MJ (2010) Diabetes: Magnitude and mechanisms. *Clin. Diabetes.* **28**:42-46

Frank C, Gonzalez MM, Oinonen C, Dunlop TW, Carlberg C (2003) Characterization of DNA complexes formed by the nuclear receptor constitutive androstane receptor *J. Biol. Chem.* **278**:43299-43310

Franks PW & Loos RJ (2006) PGC1α gene and physical activity in type 2 diabetes mellitus. *Exerc. Sport Sci. Rev.* **34**:171-175

Fukunaga BN, Probst MR, Reisz-Porszasz S & Hankinson O (1995) Identification of functional domains of the aryl hydrocarbon receptor. *J. Biol. Chem.* **270**:29270-29278

Gao J, He J, Zhai Y, Wada T & Xie W (2009) The constitutive androstane receptor is an anti-obesity nuclear receptor that improves insulin sensitivity. *J. Biol. Chem.* **284**:25984-25992

Glass CK & Rosenfeld MG (2000) The coregulator exchange in transcriptional functions of nuclear receptors. *Genes Dev.* **14**:121:141

Gronemeyer H, Gustafsson J-Å & Laudet V (2004) Principles for modulation of the nuclear receptor superfamily. *Nat. Rev. Drug. Discov.* **3**:950-964

Guarente L (2006) Sirtuins as potential targets for metabolic syndrome. *Nature.* **444**:868-874

Hankinson O (1995) The aryl hydrocarbon receptor complex. *Annu. Rev. Pharmacol. Toxicol.* **35**:307-340

Heery DM, Kalkhoven E, Hoare S & Parker MG (1997) A signature motif in transcriptional co-activators mediates binding to nuclear receptors. *Nature.* **387**:733-736

Heinzel T, Lavinsky RM, Mullen TM, Söderström M, Laherty CD, Torchia J, Yang WM, Brard G, Ngo SD, Davie JR, Seto E, Eisenman RN, Rose DW, Glass CK & Rosenfeld MG (1997) A complex containing N-CoR. mSin3 and histone deacetylase mediates transcriptional repression. *Nature.* **387**:43-48

Helsen C, Kerkhofs S, Clinckemalie L, Spans L, Laurent M, Boonen S, Vanderschueren D & Claessen F (2012) Structural basis for nuclear hormone receptor DNA binding. *Mol. Cell. Endocrinol.* **348**:411-417

Honkakoski P, Zelko I, Sueyoshi T & Negishi M (1998) The nuclear orphan receptor CAR-retinoid X receptor heterodimer activates the phenobarbital-responsive enhancer module of CYP2B gene. *Mol Cell Biol.* **18**:5652-5658

Hu X & Lazar MA (1999) The CoRNR motif controls the recruitment of corepressors by nuclear hormone receptors. *Nature.* **402**:93-96

Huss JM, Kopp RP & Kelly DP (2002) Peroxisome proliferators-activated receptor coactovator 1α (PGC1α) coactivates the cardiac-enriched nuclear receptor estrogen-related receptor-α and -γ. Identification of novel leucine-rich interaction motif within PGC1α. *J. Biol. Chem.* **277**:40265-40274

International Human Genome Sequencing Consortium (2004) Finishing the euchromatic sequence of the human genome. *Nature.* **431**:931-945

Jover R, Moya M & Gomez-Lechon MJ (2009) Transcriptional regulation of cytochrome p450 genes by nuclear receptor hepatocyte nuclear factor 4-alpha. *Curr.Drug Metab.* **10**:508-519

Kahn SE, Hull RL & Utzschneider KM (2006) Mechanisms linking obesity to insulin resistance and type 2 diabetes. *Nature.* **444**:840-846

Kelly DP & Scarpulla RC (2004) Transcriptional regulatory circuits controlling mitochondrial biogenesis and function. *Genes Dev.* **18**:357-368

Kelly T, Yang W, Chen CS, Reynolds K & He J (2008) Global burden of obesity in 2005 and projections to 2030. *Int. J. Obes. (Lond).* **32**:1431-1437

Kim B (2008) Thyroid hormone as a determinant of energy expenditure and the basal metabolic rate. *Thyroid.* **18:**141-144

Kliewer SA, Moore JT, Wade L, Staudinger JL, Watson MA, Jones SA, McKee DD, Oliver BB, Willson TM, Zetterstorm RH, Perlmann T & Lehmann JM (1998) An orphan nuclear receptor activated by pregnanes defines a novel steroid signaling pathway. *Cell.* **92**:73-82

Knutti D, Kaul A & Kralli A (2000) A tissue-specific coactivator of steroid receptors, identifies in functional genetic screen. *Mol. Cell. Biol.* **20**:2411-2422

Kodama S, Koike C, Negishi M & Yamamoto Y (2004) Nuclear receptors CAR and PXR cross talk with FOXO1 to regulate that encode drug-metabolizing and gluconeogenic enzymes. *Mol. Cell. Biol.* **24**:7931-7940

Koo SH, Satoh H, Herzig S, Lee CH, Hedrick S, Kulkarni R, Evans RM, Olefsky J & Montminy M (2004) PGC1 promotes insulin resistance in liver trough PPAR-α-dependent induction of TRB-3. *Nat. Med.* **10**:530-534

Lagauge M, Argmann C, Gerhart-Hines Z, Meziane H. Lerin C, Daussin F, Messadeg N, Milne J, Lambert P, Elliot P, Geny B, Laakso M, Puigserver P & Auwerx J (2004) Resveratrol improves mitochondrial function and protects against metabolic disease by activating SIRT1 and PGC1α. *Cell.* **127**:1109-1122

Laudet V (1997) Evolution of the nuclear receptor superfamily: early deversification from an ancectral orphan receptor. *J. Mol. Endocrinol.* **19**:207-226

Lazar MA (2003) Nuclear receptor corepressors. *Nucl. Recept. Signal.* **1**:1-4

Lefebvre P, Benomar Y & Staels B (2010) Retinoid X receptors: common heterodimerization partners with distinc functions. *Trends Endocrinol. Metab.* **21**:676-683

Lehman JJ, Barger PM, Kovacs A, Saffitz, JE, Medeiros DM & Kelly DP (2000) Peroxisome proliferator-activated receptor γ coactivator-1 promotes cardiac mitochondrial biogenesis. *J. Clin. Invest.* **106**:847-856

Lemon B & Tjian R (2000) Orchestrated response: a symphony of transcription factors for gene control. *Genes Dev.* **14**:105-178

Li X & Kazgan N (2011) Mammalian Sirtuins and Energy Metabolism. *Int. J. Biol. Sci.* **7**:575-587

Li Y, Kovach A, Suino-Powell K, Martynowski D & Xu HE (2008) Structural and biochemical basis for the binding selectivity of peroxisome proliferator-activated receptor γ to PGC1α. *J. Biol. Chem.* **283**:19132-19139

Li Y, Lambert MH & Xu HE (2003) Activation of nuclear receptors: a perspective from structural genomics. *Structure.* **11**:741-746

Maglich JM, Lobe DC & Moore JT (2009) The nuclear receptor CAR (NR1I3) regulates serum triglyceride levels under conditions of metabolic stress. *J. Lipid Res.* **50**:439-445

Maglich JM, Watson J, McMillen PJ, Goodwin B, Willson TM & Moore JT (2004) The nuclear receptor CAR is a regulator of thyroid hormone metabolism during caloric restriction. *J. Biol. Chem.* **279**:19832-19838

Maglich JM, Parks DJ, Moore LB, Collins JL, Goodwin B, Billin AN, Stoltz CA, Kliewer SA, Lambert MH, Willson TM & Moore JT (2003) Identification of a novel human constitutive androstane receptor (CAR) agonist and its use in identification of CAR target genes. *J. Biol. Chem.*

Mangelsdorf DJ, Thummel C, Beato M, Herrlich P, Schütz G, Umesono K, Blumberg B, Kastner P, Mark M, Chambon P & Evans RM (1995) The nuclear receptor superfamily: the second decade. *Cell.* **83**:835-839

McInerney EM, Rose DW, Flynn SE, Westin, S, Mullen TM, Krones A, Inostroza J, Torchia J, Nolte RT, Assa-Munt N, Milburn MV, Glass CK & Rosenfeld MG (1998) Determinant of cofactor LXXLL motif specificity in nuclear receptor transcriptional activation. *Genes Dev.***12**:3357-3368

Miao J, Fang S, Bae Y & Kemper JK (2006) Functional inhibitory cross-talk between constitutive androstane receptor and hepatic nuclear factor-4 in hepatic lipid/glucose metabolism is mediated by competition for binding to the DR1 motif and to the common coactivators, GRIP-1 and PGC1α. *J. Biol. Chem.* **281**:14537-14546

Mitra R, Nogee DP, Zechner JF, Yea K, Gierasch CM, Kovacs A, Medeiros DM, Kelly DP & Duncan JG (2012) The transcriptional coactivators, PGC1α and β, cooperate to maintain cardiac mitochondrial function during the early stages of insulin resistance. *J. Mol. Cell. Cardiol.* **52**:701-710

Molnár F, Matilainen M & Carlberg C (2005) Structural determinants of the agonist-independent association of human peroxisome proliferators-activated receptors with coactivators. *J. Biol. Chem.* **280**:26543-26556

Moore LB, Parks DJ, Jones SA, Bledsoe RK, Consler TG, Stimmel JB, Googwin B, Liddle C, Blanchard SG, Willsin TM, Collins JL & Kliewer SA (2000) Orphan nuclear receptor constitutive androstane receptor and pregnane X receptor share xenobiotic and steroid ligands. *J. Biol. Chem.* **275**:15122-15127

Moreau A, Vilarem MJ, Maurel P & Pascussi JM (2007) Xenoreceptors CAR and PXR activation and consequences on lipid metabolism, glucose homeastasism, and inflammatory response. *Mol. Pharm.* **5**:35-41

Nagy L, Kao HY, Chakravarti D, Lin RJ, Hassig Ca, Ayer DE, Schreiber SL & Evans RM (1997) Nuclear receptor repression mediated by complex containing SMRT, mSin3A, and histone deacetylase. *Cell.* **89**:373-380

Novy R, Drott D, Yaeger K & Mierendorf R (2001) Overcoming the codon bias of *E. coli* for enhanced protein expression. *Innovations.* **12**:1-3

Polonsky KS (2000) Dynamics of insulin secretion in obesity and diabetes. *Int. J. Obes. Relat. Metab. Disord.* **24**(Suppl2):S29-31

Puigserver P, Wu Z, Park CW, Graver R, Wright M & Spiegelman BM (1998) A cold-induced coactivator of nuclear receptors linked to adaptive thermogenesis. *Cell.* **92**:829-839

Quattrochi LC & Guzelian (2001) CYP3A regulation: from pharmacology to nuclear receptors. *Drug Metab. Dispos.* **29**:615-622

Rha GB, Wu G, Shoelson SE & Chi YI (2009) Multiple binding modes between HNF4α and the LXXLL motifs of PGC1alpha lead to full activation. *J. Biol. Chem.* **284**:35165-35176

Rosenfeld MG & Glass CK (2001) Coregulator codes of transcriptional regulation by nuclear receptors. *J. Biol. Chem.* **276**:36865-36868

Rosenfeld MG, Lunyak VV & Glass (2006) Sensors and signals: a coactivator/corepressor/epigenetic code for integrating signal-dependent programs of transcriptional response. *Genes Dev.* **20**:1405-1428

Sadana P & Park EA (2007) Characterization of the transactivation domain in the peroxisome-proliferator-activated receptor γ co-activator (PGC1). *Biochem. J.* **403**:511-518

Savkur RS & Burris TP (2004) The coactivator LXXLL nuclear receptor recognition motif. *J-Pept. Res.* **63**:207-212

Shiraki T, Sakai N, Kanaya E & Jingami H (2003) Activation of orphan nuclear constitutive androstane receptor requires subnuclear targeting by peroxisome proliferator-activated receptor gamma coactivator-1 alpha. A possible link between xenobiotic response and nutritional state. *J. Biol. Chem.* **278**:11344-11350

Sladek FM (2010) What are nuclear receptor ligands? *Mol. Cell. Endocrinol.* **334**:3-13

Stallcup MR, Kim JH, Teyssier C, Lee Y-H, Ma H & Chen D (2003) The roles of protein-protein interactions and protein methylation in transcriptional activation by nuclear receptors and their coactivators. *J. Steroid Biochem. Mol. Biol.* **85**:139-145

St-Pierre J, Drori S, Uldry M, Silvaggi JM, Jager S, Handschin C, Zheng K, Lin J, Yang W, Simon DK, Bachoo R & Spiegelman BM (2006) Suppression of reactive oxygen species and neurodegeneration by the PGC1 transcriptional coactivators. *Cell.* **127**:397-408

Thomas MC & Chiang C-M (2006) The general transcription machinery and general cofactors. *Crit. Rev. Biochem. Mol. Biol.* **41**:105-178

Torchia J, Glass C & Rosenfeld MG (1998) Co-activators and co-repressors in the integration of transcriptional responses. *Curr. Opin. Cell Biol.* **10**:373-383

Torok MS & Grant P (2004) Histone acetyltransferase proteins contribute to transcriptional processes at multiple levels. *Adv. Protein Chem.* **67**:181-199

Tzameli I, Pissios P, Schuetz EG, Moore DD (2000) The xenobiotic compound 1,4-bis[2-(3,5-dichloropyridyloxy)]benzene is an agonist ligand for the nuclear receptor CAR. *Mol. Cell. Biol.* **20**:2951-2958

Ueda A, Hamadeh HK, Webb HK, Yamamoto Y, Sueyoshi T, Afshari CA, Lehmann JM & Negishi M (2002) Diverse roles of the nuclear receptor CAR in regulating hepatic genes in response to phenobarbital. *Mol. Pharmacol.* **61**:1-6

Vaquerizas JM, Kummerfeld SK, Teichmann SA & Luscombe NM (2009) A census of huan transcription factors: function, expression and evolution. *Nat. Rev. Genet.* **10**:252-263

Wagner S, Klepsch M, Schlegel S, Appel A, Draheim R, Tarry M, Högbom M, van Wijk K, Slotboom D, Persson J & de Gier J-W (2008) Tuning *Escherichia coli* for membrane protein overexpression. *PNAS*. **105**:14371-14376

Wertheimer E, Gutierrez-Uzquiza A, Rosemblit C, Lopez-Haber C, Sosa MS & Kazanietz MG (2012) Rac signaling in breast cancer: a tale of GEFs and GAPs. *Cell. Signal.* **24**:353-362

Wrange Ö & Gustafsson J-Å (1978) Separation of the hormone- and DNA-binding sites of the hepatic glucocorticoid receptor by means of proteolysis. *J. Biol. Chem.* **3**:856-865

Wu Z, Puigserver P, Andersson U, Zhang C, Adelmant G, Mootha V, Troy A, Cinti S, Lowell B, Scarpulla RC & Spiegelman BM (1999) Mechanisms controlling mitochondrial biogenesis and respiration through the thermogenic coactivator PGC1. *Cell.* **98**:115-124

Wurtz J-M, Bourguet W, Renaud J-P, Vivat V, Chambon P, Moras D & Gronemeyer H (1996) A canonical structure for the ligand-binding domain of nuclear receptors. *Nat. Struct. Biol.* **3**:87-94

Xie W & Evans RM (2001) Orphan nuclear receptors: the exotics of xenobiotics. *J. Biol. Chem.* **276**:37739-37742

Xie W, Uppal H, Saini SP, Mu Y, Little JM, Radominska-Pandya A & Zemaitis MA (2004) Orpan nuclear receptor-mediated xenobiotic regulation in drug metabolism. *Drug Discov. Today.* **9**:442-449

Xie YB, Nedumaran B & Choi HS (2009) Molecular characterization of SMILE as a novel corepressor of nuclear receptors. *Nucleic Acids Res.* **37**:4100-4115

Xu L, Glass CK & Rosenfeld MG (1999) Coactivator and corepressor complexes in nuclear receptor function. *Curr. Opin. Genet. Dev.* **9**:140-147

Yamagata K, Yoshimochi K, Daitoku H, Hirota K & Fukamizu A (2007) Bile acid represses the peroxisome proliferator-activated receptor-γ coactivator-1 promoter activity in a small heterodimer partner-dependent manner. *Int. J. Mol. Med.* **19:**751-756

Yoon JC, Puigserver P, Chen G, Donovan J, Wu Z, Rhee J, Adelmant G, Stafford J, Kahn CR, Granner DK, Newgard CB & Spiegelman BM (2001) Control of hepatic gluconeogenesis through the transcriptional coactivator PGC1. *Nature.* **413**:131-138

Zelko I & Negishi M (2000) Phenobarbital-elicited activation of nuclear receptor CAR in induction of cytochrome P450 genes. *Biochem. Biophys. Res. Commun.* **277**:1-6

Zhang Y, Castellani LW, Sinal CJ, Gonzalez FJ, Edwards PA (2004) Peroxisome proliferator-activated receptor-γ coactivator 1α (PGC1α) regulates triglyceride metabolism by activation of nuclear receptor FXR. *Genes Dev.* **18**:157-169

Web sources

http://bioinf.cs.ucl.ac.uk/disopred/

http://blast.ncbi.nlm.nih.gov/

http://string-db.org/

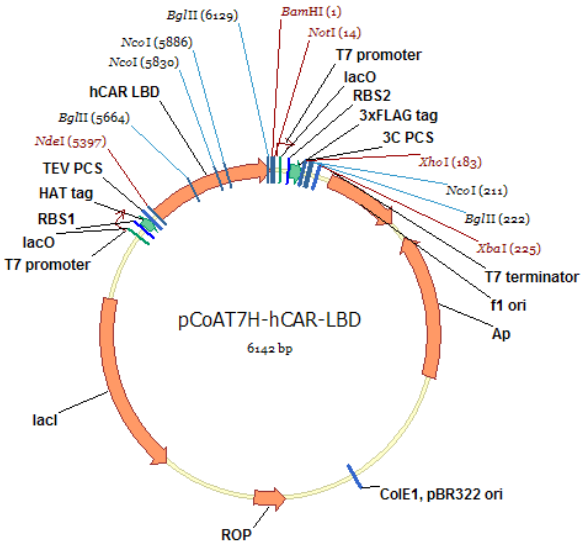
http://web.expasy.org/translate/
http://www.genscript.com/cgi-bin/tools/rare_codon_analysis
http://www.ncbi.nlm.nih.gov/nuccore
http://www.xbrane.com/files/documents/Protocol_Xbrane_Lemo_System.pdf

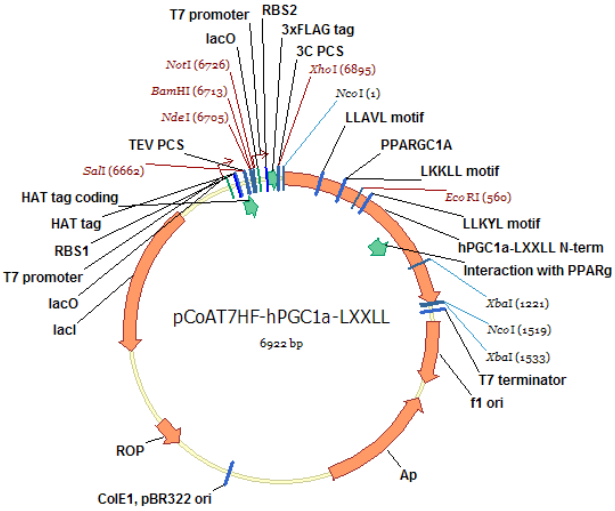
APPENDIX A

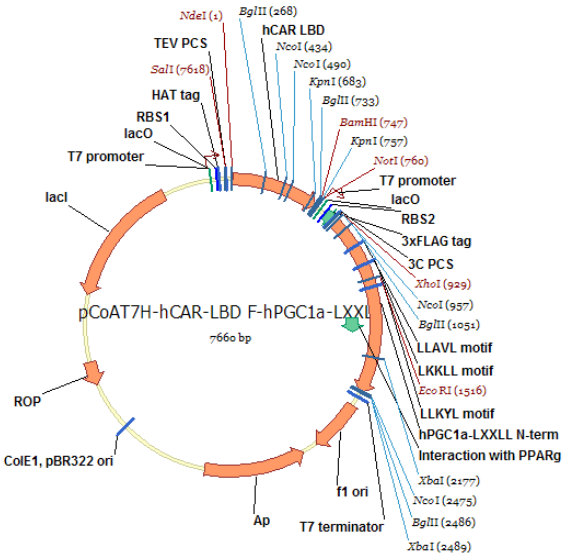
Maps of the recombinant co-expression vectors

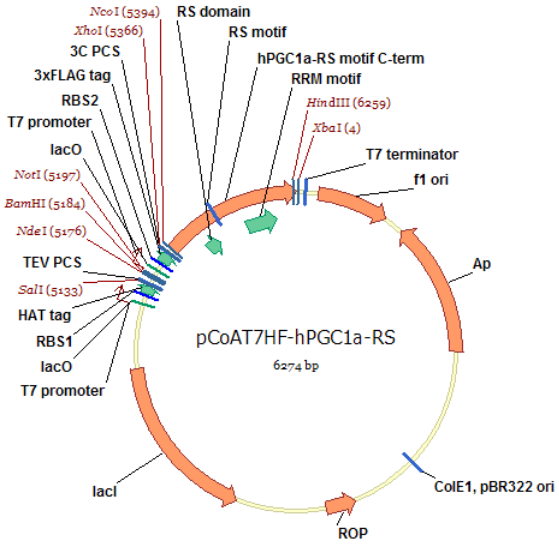
Abbreviations

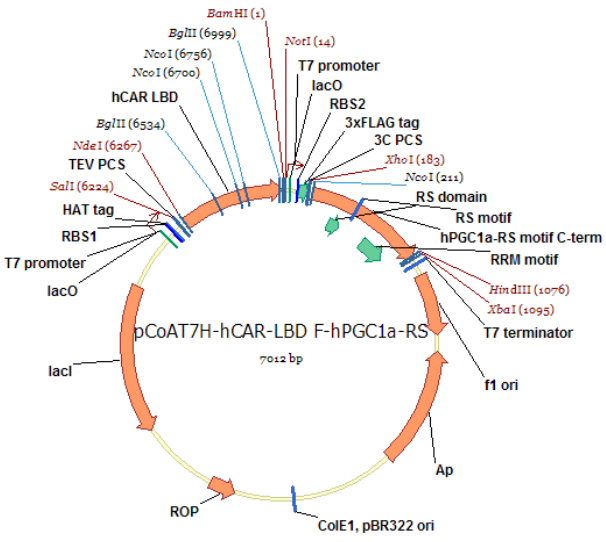
Ap	ampicillin resistance gene
3C PCS	3C protease cleavage site
TEV PCS	TEV protease cleavage site
lacI	*lac* repressor gene
lacO	*lac* operon
RBS	ribosome binding site
f1	origin of replication
ROP	repressor of primer
ColE1, pB322 ori	origin of replication

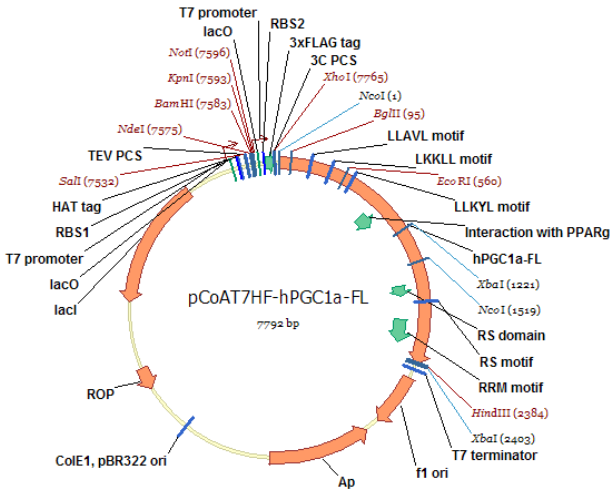


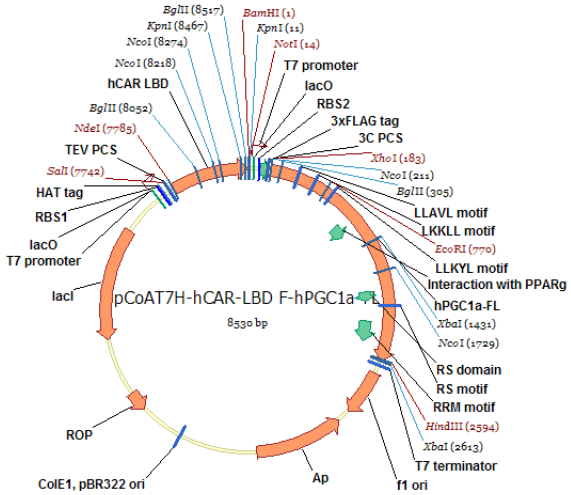





NcoI (5394)
XhoI (5366)
RS domain
RS motif
3C PCS
3xFLAG tag
hPGC1a-RS motif C-term
RRM motif
RBS2
HindIII (6259)
T7 promoter
XbaI (4)
lacO
T7 terminator
NotI (5197)
f1 ori
BamHI (5184)
NdeI (5176)
Ap
TEV PCS
SalI (5133)
HAT tag
pCoAT7HF-hPGC1a-RS
RBS1
6274 bp
lacO
T7 promoter
lacI
ColE1, pBR322 ori
ROP


BamHI (1)
BglII (6999)
NotI (14)
NcoI (6756)
T7 promoter
NcoI (6700)
lacO
hCAR LBD
RBS2
BglII (6534)
3xFLAG tag
NdeI (6267)
3C PCS
TEV PCS
XhoI (183)
SalI (6224)
NcoI (211)
HAT tag
RS domain
RBS1
RS motif
T7 promoter
hPGC1a-RS motif C-term
RRM motif
lacO
HindIII (1076)
XbaI (1095)
pCoAT7H-hCAR-LBD F-hPGC1a-RS
T7 terminator
7012 bp
lacI
f1 ori
Ap
ROP
ColE1, pBR322 ori


T7 promoter
lacO
RBS2
NotI (7596)
3xFLAG tag
3C PCS
KpnI (7593)
XhoI (7765)
BamHI (7583)
NcoI (1)
BglII (95)
NdeI (7575)
LLAVL motif
TEV PCS
LKKLL motif
SalI (7532)
EcoRI (560)
HAT tag
LLKYL motif
RBS1
Interaction with PPARg
T7 promoter
hPGC1a-FL
lacO
XbaI (1221)
lacI
pCoAT7HF-hPGC1a-FL
NcoI (1519)
7792 bp
RS domain
RS motif
ROP
RRM motif
HindIII (2384)
XbaI (2403)
ColE1, pBR322 ori
T7 terminator
f1 ori
Ap


BglII (8517)
BamHI (1)
KpnI (8467)
KpnI (11)
NcoI (8274)
NotI (14)
NcoI (8218)
T7 promoter
hCAR LBD
lacO
RBS2
BglII (8052)
3xFLAG tag
NdeI (7785)
3C PCS
TEV PCS
XhoI (183)
SalI (7742)
NcoI (211)
HAT tag
BglII (305)
RBS1
LLAVL motif
lacO
LKKLL motif
T7 promoter
EcoRI (770)
LLKYL motif
Interaction with PPARg
lacI
pCoAT7H-hCAR-LBD F-hPGC1a-FL
hPGC1a-FL
8530 bp
XbaI (1431)
NcoI (1729)
RS domain
RS motif
RRM motif
ROP
HindIII (2594)
XbaI (2613)
T7 terminator
ColE1, pBR322 ori
Ap
f1 ori

APPENDIX B

Alignment query for human full-length PGC1α (query) and mouse full-length PGC1α (subject) nucleic acid sequences

LLAVL motif (262-276), LKKLL motif (430-444), LLKYL motif (628-642), interaction with PPARγ (877-1017), RS domain (1696-1797), RS motif (1861-1900) and RRM motif (2029-2259) highlighted in grey with differences highlighted in red.

```
>lcl|10717 gi|238018130:140-2533 Mus musculus peroxisome proliferative activated receptor,
gamma, coactivator 1 alpha (Ppargc1a), transcript variant 1, mRNA

Length=2394
Score = 3175 bits (1719), Expect = 0.0
Identities = 2087/2267 (92%), Gaps = 15/2267 (1%)
Strand=Plus/Plus

Query  140   ATACAGACAGCTTTCTGGGTGGACTCAAGTGGTGCAGTGACCAATCAGAAATAATATCCA  199
             |||||||||||||||||||||||| | |||||||| || |||||||| ||||| |||||||
Sbjct  134   ATACAGACAGCTTTCTGGGTGGATTGAAGTGGTGTAGCGACCAATCGGAAATCATATCCA  193

Query  200   ATCAGTACAACAATGAGCCTTCAAACATATTTGAGAAGATAGATGAAGAGAATGAGGCAA  259
             | |||||||||||||||||| | ||||||||||||||||||||||||||||||||||||||
Sbjct  194   ACCAGTACAACAATGAGCCTGCGAACATATTTGAGAAGATAGATGAAGAGAATGAGGCAA  253

Query  260   ACTTGCTAGCAGTCCTCACAGAGACACTAGACAGTCTCCCTGTGGATGAAGACGGATTGC  319
             |||||||||| ||| ||||||||||||| ||||||||||| ||||||||||||||||||||
Sbjct  254   ACTTGCTAGCGGTTCTCACAGAGACACTGGACAGTCTCCCCGTGGATGAAGACGGATTGC  313

Query  320   CCTCATTTGATGCGCTGACAGATGGAGACGTGACCACTGACAATGAGGCTAGTCCTTCCT  379
             ||||||||||||| ||||||||||||| ||||||||||||||| ||||| ||||||||||
Sbjct  314   CCTCATTTGATGCACTGACAGATGGAGCCGTGACCACTGACAACGAGGCCAGTCCTTCCT  373

Query  380   CCATGCCTGACGGCACCCCTCCACCC-CAGGAGGCAGAAGAGCCGTCTCTACTTAAGAAG  438
             |||||||||||||||||||||||| ||| |||||||||||||||||||||||||||||||
Sbjct  374   CCATGCCTGACGGCACCCCTCC-CCCTCAGGAGGCAGAAGAGCCGTCTCTACTTAAGAAG  432

Query  439   CTCTTACTGGCACCAGCCAACACTCAGCTAAGTTATAATGAATGCAGTGGTCTCAGTACC  498
             ||||||||||||||||||||||||||| || || ||||||||||| ||||| || ||
Sbjct  433   CTCTTACTGGCACCAGCCAACACTCAGCTCAGCTACAATGAATGCAGCGGTCTTAGCACT  492

Query  499   CAGAACCAT---GCAAATCACAATCACAGGATCAGAACAAACCCTGCAATTGTTAAGACT  555
             |||||||||   ||||| ||||  |||||||||||||||||||||||| ||||||||||
Sbjct  493   CAGAACCATGCAGCAAACCACACCCACAGGATCAGAACAAACCCTGCCATTGTTAAGACC  552

Query  556   GAGAATTCATGGAGCAATAAAGCGAAGAGTATTTGTCAACAGCAAAAGCCACAAAGACGT  615
             |||||||||||||||||||||||||||||| ||||||||||||||||||||||||||||||
Sbjct  553   GAGAATTCATGGAGCAATAAAGCGAAGAGCATTTGTCAACAGCAAAAGCCACAAAGACGT  612

Query  616   CCCTGCTCGGAGCTTCTCAAATATCTGACCACAAACGATGACCCTCCTCACACCAAACCC  675
             |||||||| |||||||||||| |||||||||||||||||||||||||||||||||||||||
Sbjct  613   CCCTGCTCAGAGCTTCTCAAGTATCTGACCACAAACGATGACCCTCCTCACACCAAACCC  672

Query  676   ACAGAGAACAGAAACAGCAGCAGAGACAAATGCACCTCCAAAAAGAAGTCCCACACACAG  735
             ||||| ||||| |||||||||||||||||||  | || ||||||||||||||| |||||
Sbjct  673   ACAGAAAACAGGAACAGCAGCAGAGACAAATGTGCTTCGAAAAAGAAGTCCCATACACAA  732

Query  736   TCGCAGTCACAACACT--TACAAGCCAAACCAACAACTTTATCTCTTCCTCTGACCCCAG  793
              ||||||| ||||| |   | ||||||||||||||||||||||||||||||||||||||||
Sbjct  733   CCGCAGTCGCAACA-TGCT-CAAGCCAAACCAACAACTTTATCTCTTCCTCTGACCCCAG  790

Query  794   AGTCACCAAATGACCCCAAGGGTTCCCCATTTGAGAACAAGACTATTGAACGCACCTTAA  853
             |||||||||||||||||||||||||||||||||||||||||||||||||||| || |||||||
Sbjct  791   AGTCACCAAATGACCCCAAGGGTTCCCCATTTGAGAACAAGACTATTGAGCGAACCTTAA  850

Query  854   GTGTGGAACTCTCTGGAACTGCAGGCCTAACTCCACCCACCACTCCTCCTCATAAAGCCA  913
             |||||||||||||||||||||||||||||||||| |||||| |||||||||||||||||||
Sbjct  851   GTGTGGAACTCTCTGGAACTGCAGGCCTAACTCCTCCCACAACTCCTCCTCATAAAGCCA  910
```

```
Query  914   ACCAAGATAACCCTTTTAGGGCTTCTCCAAAGCTGAAGTCCTCTTGCAAGACTGTGGTGC  973
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  911   ACCAAGATAACCCTTTCAAGGCTTCGCCAAAGCTGAAGCCCTCTTGCAAGACCGTGGTGC  970

Query  974   CACCACCATCAAAGAAG-CCCAGGTACAGTGAGTCTTCTGGTACACAAGGCAATAACTCC  1032
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  971   CACCGCCAACCAAGAGGGCCC-GGTACAGTGAGTGTTCTGGTACCCAAGGCAGCCACTCC  1029

Query  1033  ACCAAGAAAGGGCCGGAGCAATCCGAGTTGTATGCACAACTCAGCAAGTCCTCAGTCCTC  1092
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1030  ACCAAGAAAGGGCCCGAGCAATCTGAGTTGTACGCACAACTCAGCAAGTCCTCAGGGCTC  1089

Query  1093  A-CTGGTGGACACGAGGAAAGGAAGACCAAGCGGCCCAGTCTGCGGCTGTTTGGTGACCA  1151
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1090  AGCCGA-GGACACGAGGAAAGGAAGACTAAACGGCCCAGTCTCCGGCTGTTTGGTGACCA  1148

Query  1152  TGACTATTGCCAGTCAATTAATTCCAAAACAGAAATACTCATTAATATATCACAGGAGCT  1211
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1149  TGACTACTGTCAGTCACTCAATTCCAAAACGGATATACTCATTAACATATCACAGGAGCT  1208

Query  1212  CCAAGACTCTAGACAACTAGAAAAT-AAAGATGTCTCCTCTGATTGGCAGGGGCAGATTT  1270
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1209  CCAAGACTCTAGACAACTAGACT-TCAAAGATGCCTCCTGTGACTGGCAGGGGCACATCT  1267

Query  1271  GTTCTTCCACAGATTCAGACCAGTGCTACCTGAGAGAGACTTTGGAGGCAAGCAAGCAGG  1330
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1268  GTTCTTCCACAGATTCAGGCCAGTGCTACCTGAGAGAGACTTTGGAGGCCAGCAAGCAGG  1327

Query  1331  TCTCTCCTTGCAGCACAAGAAAACAGCTCCAAGACCAGGAAATCCGAGCCGAGCTGAACA  1390
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1328  TCTCTCCTTGCAGCACCAGAAAACAGCTCCAAGACCAGGAAATCCGAGCGGAGCTGAACA  1387

Query  1391  AGCACTTCGGTCATCCCAGTCAAGCTGTTTTTGACGACGAAGCAGACAAGACCGGTGAAC  1450
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1388  AGCACTTCGGTCATCCCTGTCAAGCTGTGTTTGACGACAAATCAGACAAGACCAGTGAAC  1447

Query  1451  TGAGGGACAGTGATTTCAGTAATGAACAATTCTCCAAACTACCTATGTTTATAAATTCAG  1510
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1448  TAAGGGATGGCGACTTCAGTAATGAACAATTCTCCAAACTACCTGTGTTTATAAATTCAG  1507

Query  1511  GACTAGCCATGGATGGCCTGTTTGATGACAGCGAAGATGAAAGTGATAAACTGAGCTACC  1570
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1508  GACTAGCCATGGATGGCCTATTTGATGACAGTGAAGATGAAAGTGATAAACTGAGCTACC  1567

Query  1571  CTTGGGATGGCACGCAATCCTATTCATTGTTCAATGTGTCTCCTTCTTGTTCTTCTTTTA  1630
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1568  CTTGGGATGGCACGCAGCCCTATTCATTGTTCGATGTGTCGCCTTCTTGCTCTTCCTTTA  1627

Query  1631  ACTCTCCATGTAGAGATTCTGTGTCACCACCCAAATCCTTATTTTCTCAAAGACCCCAAA  1690
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1628  ACTCTCCGTGTCGAGACTCAGTGTCACCACCGAAATCCTTATTTTCTCAAAGACCCCAAA  1687

Query  1691  GGATGCGCTCTCGTTCAAGGTCCTTTTCTCGACACAGGTCGTGTTCCCGATCACCATATT  1750
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1688  GGATGCGCTCTCGTTCAAGATCCTTTTCTCGACACAGGTCGTGTTCCCGATCACCATATT  1747

Query  1751  CCAGGTCAAGATCAAGGTCTCCAGGCAGTAGATCCTCTTCAAGATCCTGCTATTACTATG  1810
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1748  CCAGGTCAAGATCAAGGTCCCCAGGCAGTAGATCCTCTTCAAGATCCTGTTACTACTATG  1807

Query  1811  AGTCAAGCCACTACAGACACCGCACGCACCGAAATTCTCCCTTGTATGTGAGATCACGTT  1870
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1808  AATCAAGCCACTACAGACACCGCACACACCGCAATTCTCCCTTGTATGTGAGATCACGTT  1867

Query  1871  CAAGATCGCCCTACAGCCGTCGGCCCAGGTATGACAGCTACGAGGAATATCAGCACGAGA  1930
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1868  CAAGGTCACCCTACAGCCGTAGGCCCAGGTACGACAGCTATGAAGCCTATGAGCACGAAA  1927

Query  1931  GGCTGAAGAGGGAAGAATATCGCAGAGAGTATGAGAAGCGAGAGTCTGAGAGGGCCAAGC  1990
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1928  GGCTCAAGAGGGATGAATACCGCAAAGAGCACGAGAAGCGGGAGTCTGAAAGGGCCAAAC  1987

Query  1991  AAAGGGAGAGGCAGAGGCAGAAGGCAATTGAAGAGCGCCGTGTGATTTATGTCGGTAAAA  2050
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  1988  AGAGAGAGAGGCAGAAGCAGAAAGCAATTGAAGAGCGCCGTGTGATTTACGTTGGTAAAA  2047
```

```
Query  2051  TCAGACCTGACACAACACGGACAGAACTGAGGGACCGTTTTGAAGTTTTTGGTGAAATTG  2110
             |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2048  TCAGACCTGACACAACGCGGACAGAATTGAGAGACCGCTTTGAAGTTTTTGGTGAAATTG  2107

Query  2111  AGGAGTGCACAGTAAATCTGCGGGATGATGGAGACAGCTATGGTTTCATTACCTACCGTT  2170
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2108  AGGAATGCACCGTAAATCTGCGGGATGATGGAGACAGCTATGGTTTCATGACCTACCGTT  2167

Query  2171  ATACCTGTGATGCTTTTGCTGCTCTTGAAAATGGATACACTTTGCGCAGGTCAAACGAAA  2230
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2168  ACACCTGTGACGCTTTCGCTGCTCTTGAGAATGGATATACTTTACGCAGGTCGAACGAAA  2227

Query  2231  CTGACTTTGAGCTGTACTTTTGTGGACGCAAGCAATTTTTCAAGTCTAACTATGCAGACC  2290
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Sbjct  2228  CTGACTTCGAGCTGTACTTTTGTGGACGGAAGCAATTTTTCAAGTCTAACTATGCAGACC  2287

Query  2291  TAGATTCAAACTCAGATGACTTTGACCCTGCTTCCACCAAGAGCAAGTATGACTCTCTGG  2350
             ||||| | |||||||| || ||||||||||||||||||||||||||||||||||||||||||
Sbjct  2288  TAGATACCAACTCAGACGATTTTGACCCTGCTTCCACCAAGAGCAAGTATGACTCTCTGG  2347

Query  2351  ATTTTGATAGTTTACTGAAAGAAGCTCAGAGAAGCTTGCGCAGGTAA  2397
             |||||||||||||||||||| |||||||||||||||||||||||||||
Sbjct  2348  ATTTTGATAGTTTACTGAAGGAAGCTCAGAGAAGCTTGCGCAGGTAA  2394
```

Alignment query for human full-length PGC1α (query) and mouse full-length PGC1α (subject) amino acid sequences

LLAVL motif (88-92), LKKLL motif (144-148), LLKYL motif (210-214), interaction with PPARγ (293-338), RS domain (566-599), RS motif (621-634) and RRM (677-754) motif highlighted in grey and differences in red.

```
>lcl|6151 unnamed protein product
Length=797

 Score = 1402 bits (3628),  Expect = 0.0, Method: Compositional matrix adjust.
 Identities = 755/799 (94%), Positives = 776/799 (97%), Gaps = 3/799 (0%)

Query  1    MAWDMCNQDSESVWSDIECAALVGEDQPLCPDLPELDLSELDVNDLDTDSFLGGLKWCSD  60
            MAWDMC+QDS  VWSDIECAALVGEDQPLCPDLPELDLSELDVNDLDTDSFLGGLKWCSD
Sbjct  1    MAWDMCSQDS--VWSDIECAALVGEDQPLCPDLPELDLSELDVNDLDTDSFLGGLKWCSD  58

Query  61   QSEIISNQYNNEPSNIFEKIDEENEANLLAVLTETLDSLPVDEDGLPSFDALTDGDVTTD  120
            QSEIISNQYNNEP+NIFEKIDEENEANLLAVLTETLDSLPVDEDGLPSFDALTDG VTTD
Sbjct  59   QSEIISNQYNNEPANIFEKIDEENEANLLAVLTETLDSLPVDEDGLPSFDALTDGAVTTD  118

Query  121  NEASPSSMPDGTPPPQEAEEPSLLKKLLLAPANTQLSYNECSGLSTQNHA-NHNHRIRTN  179
            NEASPSSMPDGTPPPQEAEEPSLLKKLLLAPANTQLSYNECSGLSTQNHA NH HRIRTN
Sbjct  119  NEASPSSMPDGTPPPQEAEEPSLLKKLLLAPANTQLSYNECSGLSTQNHAANHTHRIRTN  178

Query  180  PAIVKTENSWSNKAKSICQQQKPQRRPCSELLKYLTTNDDPPHTKPTENRNSSRDKCTSK  239
            PAIVKTENSWSNKAKSICQQQKPQRRPCSELLKYLTTNDDPPHTKPTENRNSSRDKC SK
Sbjct  179  PAIVKTENSWSNKAKSICQQQKPQRRPCSELLKYLTTNDDPPHTKPTENRNSSRDKCASK  238

Query  240  KKSHTQSQSQHLQAKPTTLSLPLTPESPNDPKGSPFENKTIERTLSVELSGTAGLTPPTT  299
            KKSHTQ QSQH QAKPTTLSLPLTPESPNDPKGSPFENKTIERTLSVELSGTAGLTPPTT
Sbjct  239  KKSHTQPQSQHAQAKPTTLSLPLTPESPNDPKGSPFENKTIERTLSVELSGTAGLTPPTT  298

Query  300  PPHKANQDNPFRASPKLKSSCKTVVPPPSKKPRYSESSGTQGNNSTKKGPEQSELYAQLS  359
            PPHKANQDNPF+ASPKLK SCKTVVPPP+K+ RYSE SGTQG++STKKGPEQSELYAQLS
Sbjct  299  PPHKANQDNPFKASPKLKPSCKTVVPPPTKRARYSECSGTQGSHSTKKGPEQSELYAQLS  358

Query  360  KSSVLTGGHEERKTKRPSLRLFGDHDYCQSINSKTLILINISQELQDSRQLENKDVSSDW  419
            KSS L+ GHEERKTKRPSLRLFGDHDYCQS+NSKT+ILINISQELQDSRQL+ KD S DW
Sbjct  359  KSSGLSRGHEERKTKRPSLRLFGDHDYCQSLNSKTDILINISQELQDSRQLDFKDASCDW  418

Query  420  QGQICSSTDSDQCYLRETLEASKQVSPCSTRKQLQDQEIRAELNKHFGHPSQAVFDDEAD  479
            QG ICSSTDS QCYLRETLEASKQVSPCSTRKQLQDQEIRAELNKHFGHP QAVFDD++D
Sbjct  419  QGHICSSTDSGQCYLRETLEASKQVSPCSTRKQLQDQEIRAELNKHFGHPGQAVFDDKSD  478
```

```
Query  480  KTQELRDSDFSNEQFSKLPMFINSGLAMDGLFDDSEDESDKLSYPWDGTQSYSLFNVSPS  539
            KT EELRD DFSNEQFSKLP+FINSGLAMDGLFDDSEDESDKLSYPWDGTQ YSLF+VSPS
Sbjct  479  KTSELRDGDFSNEQFSKLPVFINSGLAMDGLFDDSEDESDKLSYPWDGTQFYSLFDVSPS  538

Query  540  CSSFNSPCRDSVSPPKSLFSQRPQRMRSRSRSFSRHRSCSRSPYSRSRSRSPGSRSSSRS  599
            CSSFNSPCRDSVSPPKSLFSQRPQRMRSRSRSFSRHRSCSRSPYSRSRSRSPGSRSSSRS
Sbjct  539  CSSFNSPCRDSVSPPKSLFSQRPQRMRSRSRSFSRHRSCSRSPYSRSRSRSPGSRSSSRS  598

Query  600  CYYYESSHYRHRTHRNSPLYVRSRSRSPYSRRPRYDSYEEYQHERLKREEYRREYEKRES  659
            CYYYESSHYRHRTHRNSPLYVRSRSRSPYSRRPRYDSYE Y+HERLKR+EYR+E+EKRES
Sbjct  599  CYYYESSHYRHRTHRNSPLYVRSRSRSPYSRRPRYDSYEAYEHERLKRDEYRKEHEKRES  658

Query  660  ERAKQRERQRQKAIEERRVIYVGKIRPDTTRTELRDRFEVFGEIEECTVNLRDDGDSYGF  719
            ERAKQRERQ+QKAIEERRVIYVGKIRPDTTRTELRDRFEVFGEIEECTVNLRDDGDSYGF
Sbjct  659  ERAKQRERQKQKAIEERRVIYVGKIRPDTTRTELRDRFEVFGEIEECTVNLRDDGDSYGF  718

Query  720  ITYRYTCDAFAALENGYTLRRSNETDFELYFCGRKQFFKSNYADLDSNSDDFDPASTKSK  779
            ITYRYTCDAFAALENGYTLRRSNETDFELYFCGRKQFFKSNYADLD+NSDDFDPASTKSK
Sbjct  719  ITYRYTCDAFAALENGYTLRRSNETDFELYFCGRKQFFKSNYADLDTNSDDFDPASTKSK  778

Query  780  YDSLDFDSLLKEAQRSLRR  798
            YDSLDFDSLLKEAQRSLRR
Sbjct  779  YDSLDFDSLLKEAQRSLRR  797
```

Alignment query for human PGC1α-LXXLL (query) and mouse PGC1α-LXXLL (subject) amino acid sequences

LLAVL motif (88-92), LKKLL motif (144-148), LLKYL motif (210-214) and interaction with PPARγ (293-338) highlighted in grey and differences in red.

```
>lcl|61693 unnamed protein product

Length=506
Score =  946 bits (2444), Expect = 0.0, Method: Compositional matrix adjust.
Identities = 473/508 (93%), Positives = 487/508 (96%), Gaps = 3/508 (1%)

Query  1    MAWDMCNQDSESVWSDIECAALVGEDQPLCPDLPELDLSELDVNDLDTDSFLGGLKWCSD  60
            MAWDMC+QDS  VWSDIECAALVGEDQPLCPDLPELDLSELDVNDLDTDSFLGGLKWCSD
Sbjct  1    MAWDMCSQDS--VWSDIECAALVGEDQPLCPDLPELDLSELDVNDLDTDSFLGGLKWCSD  58

Query  61   QSEIISNQYNNEPSNIFEKIDEENEANLLAVLTETLDSLPVDEDGLPSFDALTDGDVTTD  120
            QSEIISNQYNNEP+NIFEKIDEENEANLLAVLTETLDSLPVDEDGLPSFDALTDG VTTD
Sbjct  59   QSEIISNQYNNEPANIFEKIDEENEANLLAVLTETLDSLPVDEDGLPSFDALTDGAVTTD  118

Query  121  NEASPSSMPDGTPPPQEAEEPSLLKKLLLAPANTQLSYNECSGLSTQNHA-NHNHRIRTN  179
            NEASPSSMPDGTPPPQEAEEPSLLKKLLLAPANTQLSYNECSGLSTQNHA NH HRIRTN
Sbjct  119  NEASPSSMPDGTPPPQEAEEPSLLKKLLLAPANTQLSYNECSGLSTQNHAANHTHRIRTN  178

Query  180  PAIVKTENSWSNKAKSICQQQKPQRRPCSELLKYLTTNDDPPHTKPTENRNSSRDKCTSK  239
            PAIVKTENSWSNKAKSICQQQKPQRRPCSELLKYLTTNDDPPHTKPTENRNSSRDKC SK
Sbjct  179  PAIVKTENSWSNKAKSICQQQKPQRRPCSELLKYLTTNDDPPHTKPTENRNSSRDKCASK  238

Query  240  KKSHTQSQSQHLQAKPTTLSLPLTPESPNDPKGSPFENKTIERTLSVELSGTAGLTPPTT  299
            KKSHTQ QSQH QAKPTTLSLPLTPESPNDPKGSPFENKTIERTLSVELSGTAGLTPPTT
Sbjct  239  KKSHTQPQSQHAQAKPTTLSLPLTPESPNDPKGSPFENKTIERTLSVELSGTAGLTPPTT  298

Query  300  PPHKANQDNPFRASPKLKSSCKTVVPPPSKKPRYSESSGTQGNNSTKKGPEQSELYAQLS  359
            PPHKANQDNPF+ASPKLK SCKTVVPPP+K+ RYSE SGTQG++STKKGPEQSELYAQLS
Sbjct  299  PPHKANQDNPFKASPKLKPSCKTVVPPPTKRARYSECSGTQGSHSTKKGPEQSELYAQLS  358

Query  360  KSSVLTGGHEERKTKRPSLRLFGDHDYCQSINSKTEILINISQELQDSRQLENKDVSSDW  419
            KSS L+ GHEERKTKRPSLRLFGDHDYCQS+NSKT+ILINISQELQDSRQL+ KD S DW
Sbjct  359  KSSGLSRGHEERKTKRPSLRLFGDHDYCQSLNSKTDILINISQELQDSRQLDFKDASGDW  418

Query  420  QGQICSSTDSDQCYLRETLEASKQVSPCSTRKQLQDQEIRAELNKHFGHPSQAVFDDEAD  479
            QG ICSSTDS QCYLRETLEASKQVSPCSTRKQLQDQEIRAELNKHFGHP QAVFDD++D
Sbjct  419  QGHICSSTDSGQCYLRETLEASKQVSPCSTRKQLQDQEIRAELNKHFGHPGQAVFDDKSD  478
```

```
Query  480  KTGELRDSDFSNEQFSKLPMFINSGLAM  507
            KT ELRD DFSNEQFSKLP+FINSGLAM
Sbjct  479  KTSELRDGDFSNEQFSKLPVFINSGLAM  506
```

Alignment query for human PGC1α-RS (query) and mouse PGC1α-RS (subject) amino acid sequences

RS domain (60-93), RS motif (115-127) and RRM (171-247) motif highlighted in grey and differences highlighted in red.

```
>lcl|34709 unnamed protein product

Length=291
Score =  442 bits (1136), Expect = 5e-129, Method: Compositional matrix adjust.
Identities = 282/291 (97%), Positives = 289/291 (99%), Gaps = 0/291 (0%)

Query  2    DGLFDDSEDESDKLSYPWDGTQSYSLFNVSPSCSSFNSPCRDSVSPPKSLFSQRPQRMRS  61
            DGLFDDSEDESDKLSYPWDGTQ YSLF+VSPSCSSFNSPCRDSVSPPKSLFSQRPQRMRS
Sbjct  1    DGLFDDSEDESDKLSYPWDGTQPYSLFDVSPSCSSFNSPCRDSVSPPKSLFSQRPQRMRS  60

Query  62   RSRSFSRHRSCSRSPYSRSRSRSPGSRSSSRSCYYYESSHYRHRTHRNSPLYVRSRSRSP  121
            RSRSFSRHRSCSRSPYSRSRSRSPGSRSSSRSCYYYESSHYRHRTHRNSPLYVRSRSRSP
Sbjct  61   RSRSFSRHRSCSRSPYSRSRSRSPGSRSSSRSCYYYESSHYRHRTHRNSPLYVRSRSRSP  120

Query  122  YSRRPRYDSYEEYQHERLKREEYRREIEKRESERAKQRERQRQKAIEERRVIYVGKIRPD  181
            YSRRPRYDSYE Y+HERLKR+EYR+E+EKRESERAKQRERQ+QKAIEERRVIYVGKIRPD
Sbjct  121  YSRRPRYDSYEAYEHERLKRDEYRKEHEKRESERAKQRERQKQKAIEERRVIYVGKIRPD  180

Query  182  TTRTELRDRFEVFGEIEECTVNLRDDGDSYGFITYRYTCDAFAALENGYTLRRSNETDFE  241
            TTRTELRDRFEVFGEIEECTVNLRDDGDSYGFITYRYTCDAFAALENGYTLRRSNETDFE
Sbjct  181  TTRTELRDRFEVFGEIEECTVNLRDDGDSYGFITYRYTCDAFAALENGYTLRRSNETDFE  240

Query  242  LYFCGRKQFFKSNYADLDSNSDDFDPASTKSKYDSLDFDSLLKEAQRSLRR  292
            LYFCGRKQFFKSNYADLD+NSDDFDPASTKSKYDSLDFDSLLKEAQRSLRR
Sbjct  241  LYFCGRKQFFKSNYADLDTNSDDFDPASTKSKYDSLDFDSLLKEAQRSLRR  291
```